



**UNIVERSIDADE FEDERAL DE GOIÁS**  
**FACULDADE DE CIÊNCIAS E TECNOLOGIA**  
**CURSO DE GEOLOGIA**

**LUCAS MARQUES BARROS**

**APLICAÇÃO DE MACHINE LEARNING PARA IDENTIFICAÇÃO DE CORPOS  
MÁFICOS-ULTRAMÁFICOS NA PORÇÃO SUL DA FAIXA BRASÍLIA, GOIÁS,  
BRASIL**

**APARECIDA DE GOIÂNIA**  
**AGOSTO/2024**



UNIVERSIDADE FEDERAL DE GOIÁS  
FACULDADE DE CIÊNCIAS E TECNOLOGIA

## TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TRABALHO DE CONCLUSÃO DE CURSO DE GRADUAÇÃO NO REPOSITÓRIO INSTITUCIONAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio do Repositório Institucional (RI/UFG), regulamentado pela Resolução CEPEC no 1240/2014, sem ressarcimento dos direitos autorais, de acordo com a Lei no 9.610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo dos Trabalhos de Conclusão dos Cursos de Graduação disponibilizado no RI/UFG é de responsabilidade exclusiva dos autores. Ao encaminhar(em) o produto final, o(s) autor(a)(es)(as) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

### 1. Identificação do Trabalho de Conclusão de Curso de Graduação (TCCG)

Nome(s) completo(s) do(a)(s) autor(a)(es)(as): Lucas Marques Barros

Título do trabalho: Aplicação de Machine Learning para Identificação de Corpos Máficos-Ultramáficos na Porção Sul da Faixa Brasília, Goiás, Brasil

### 2. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador) Concorda com a liberação total do documento [ X ] SIM [ ] NÃO<sup>1</sup>

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante: a) consulta ao(à)(s) autor(a)(es)(as) e ao(à) orientador(a); b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo do TCCG. O documento não será disponibilizado durante o período de embargo.

#### Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro.

**Obs.: Este termo deve ser assinado no SEI pelo orientador e pelo autor.**



Documento assinado eletronicamente por **Filipe Altoe Temporim, Professor do Magistério Superior**, em 04/08/2024, às 19:36, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Lucas Marques Barros, Discente**, em 05/08/2024, às 12:07, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site [https://sei.ufg.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **4716714** e o código CRC **32B064D6**.

---

**Referência:** Processo nº 23070.038226/2024-31

SEI nº 4716714

**LUCAS MARQUES BARROS**

**APLICAÇÃO DE MACHINE LEARNING PARA IDENTIFICAÇÃO DE CORPOS  
MÁFICOS-ULTRAMÁFICOS NA PORÇÃO SUL DA FAIXA BRASÍLIA, GOIÁS,  
BRASIL**

**Trabalho de Conclusão de Curso de  
Graduação apresentado ao Curso de Geologia  
da Universidade Federal de Goiás como  
requisito para obtenção do grau de Bacharel  
em Geologia.**

**Orientador:** Prof. Dr. Filipe Altoé Temporim  
**Co-orientador:** Geól. MSc. Luiz Fernando  
Cursino Tabosa

**APARECIDA DE GOIÂNIA**

**AGOSTO/2024**

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Barros, Lucas Marques

Aplicação de Machine Learning para Identificação de Corpos Máficos- Ultramáficos na Porção Sul da Faixa Brasília, Goiás, Brasil [manuscrito] / Lucas Marques Barros. - 2024.

xvii, 72 f.: il.

Orientador: Prof. Dr. Filipe Altoé Temporim; co-orientador Luiz Fernando Cursino Tabosa.

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Goiás, Campus Aparecida de Goiânia, Geologia, Aparecida de Goiânia, 2024.

Bibliografia. Apêndice.

Inclui siglas, mapas, fotografias, abreviaturas, símbolos, tabelas, lista de figuras, lista de tabelas.

1. Machine Learning. 2. Corpos Máficos-Ultramáficos. 3. Complexo Anápolis-Itauçu. 4. Sequência Anicuns-Itaberaí. 5. Algoritmo Random Forest. I. Temporim, Filipe Altoé, orient. II. Título.

CDU 55



UNIVERSIDADE FEDERAL DE GOIÁS  
FACULDADE DE CIÊNCIAS E TECNOLOGIA

## ATA DE DEFESA DE TRABALHO DE CONCLUSÃO DE CURSO

Aos 2 dias do mês de agosto do ano de 2024 iniciou-se a sessão pública de defesa do Trabalho de Conclusão de Curso (TCC) intitulado “Aplicação de Machine Learning para Identificação de Corpos Máficos-Ultramáficos na Porção Sul da Faixa Brasília, Goiás, Brasil”, de autoria de Lucas Marques Barros, do curso de geologia, da Faculdade de Ciências e Tecnologia da UFG. Os trabalhos foram instalados pelo Dr. Filipe Altoé Temporim (FCT/UFG) com a participação dos demais membros da Banca Examinadora: Elder Yokoyama (IG/UnB) e Débora Nogueira Lopes (FCT/UFG). Após a apresentação, a banca examinadora realizou a arguição do estudante. Posteriormente, de forma reservada, a Banca Examinadora atribuiu a nota final de 9.5, tendo sido o TCC considerado aprovado. Proclamados os resultados, os trabalhos foram encerrados e, para constar, lavrou-se a presente ata que segue assinada pelos Membros da Banca Examinadora.



Documento assinado eletronicamente por **Debora Nogueira Lopes, Professor do Magistério Superior**, em 03/08/2024, às 07:00, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Filipe Altoe Temporim, Professor do Magistério Superior**, em 04/08/2024, às 19:35, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site [https://sei.ufg.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **4716711** e o código CRC **F6100E66**.

*“Dedico este trabalho aos meus amados familiares, cujo apoio inabalável e incentivo constante foram fundamentais em minha jornada acadêmica e pessoal. A vocês, expresso minha profunda gratidão.”*

## AGRADECIMENTOS

A jornada acadêmica é, por essência, um caminho repleto de desafios, superações e aprendizados. No decorrer deste percurso, inúmeras pessoas contribuíram de maneira significativa para que eu alcançasse este momento de conclusão. Este trabalho é fruto de um esforço coletivo e gostaria de expressar minha mais profunda gratidão a todos que, direta ou indiretamente, estiveram ao meu lado.

Primeiramente, agradeço à minha família, minha base e sustento inabalável. À minha mãe, Maria Angélica, e ao meu pai, Alson, por seu amor incondicional, paciência e apoio irrestrito. Vocês foram o alicerce que me manteve firme, mesmo nos momentos mais desafiadores. À minha irmã, Rafaella, por ser uma constante fonte de inspiração e encorajamento.

Aos meus amigos, cujas presenças foram fundamentais para a minha trajetória, expresso meu sincero agradecimento. Em especial, a Lucas, Paulo Ricardo, Paulo, Flávio, Gustavo, Roberto, Dedé, Ícaro, Juca, Débora, Vilson, Ivanir, Nathalia, Yuri, Geovane, Gabriel, Vinícius e Brunna. Vocês construíram o meu caráter e foram minhas âncoras, oferecendo suporte emocional e momentos de descontração e alegria que aliviaram o peso da caminhada acadêmica.

Às turmas de 2016, 2017 e 2018, companheiros de batalha que compartilharam comigo as angústias e vitórias ao longo dos anos, deixo aqui minha gratidão. Juntos, construímos uma rede de apoio e amizade que transcende as salas de aula.

Ao corpo técnico, cujos esforços muitas vezes silenciosos foram cruciais para o sucesso desta jornada, deixo minha gratidão. Mateus, Marcília, Ivan, Cláudia, Francisco, Bruno e Karen, agradeço pelo suporte e paciência.

Agradeço também aos professores e professoras do corpo docente do curso de geologia, em especial ao meu orientador, o senhor Prof. Dr. Filipe Temporim, por compartilhar seu vasto conhecimento. Sua liderança foi essencial para a concretização deste trabalho.

Por fim, agradeço aos projetos Magnetismo do Magmatismo Tardio da Faixa Brasília e Paleogeografia do Gondwana Ocidental; e Evolução geológica e potencial metalogenético da região do contato entre o núcleo metamórfico da Faixa Brasília e o Arco Magmático de Arenópolis; pelo financiamento, que possibilitou a realização desta pesquisa.

A todos, meu muito obrigado.

- Lucas Marques Barros

*“Nada na vida deve ser temido, somente compreendido. Agora é hora de compreender mais para temer menos.”*

- Marie Curie

## RESUMO

Este trabalho investiga a aplicação de *Machine Learning* na identificação de corpos máficos-ultramáficos em uma área de 200 km<sup>2</sup> na região sul da Faixa Brasília, focando especificamente no limite entre a sequência Anicuns-Itaberaí e o Complexo Anápolis-Itauçu, nas proximidades do município de Anicuns, Goiás. O estudo empregou o algoritmo *Random Forest* para processar e analisar dados geofísicos e de sensoriamento remoto, com o objetivo de aprimorar o mapeamento geológico e a detecção desses corpos. A pesquisa integrou dados aerogeofísicos, incluindo magnetometria e radiometria, e dados de sensoriamento remoto multiespectrais e de radar para criar um robusto conjunto de dados geoespaciais. O modelo foi treinado com uma fração significativa da base de dados para produzir um mapa preditivo das ocorrências geológicas. Os resultados obtidos demonstraram uma alta eficácia do modelo, com precisão significativa na discriminação dos corpos geológicos de interesse, validados por comparações com dados de campo e proporcionando uma nova perspectiva sobre a distribuição e características dos corpos máficos-ultramáficos na área estudada. Este estudo evidencia o potencial do uso de técnicas de *Machine Learning* na geologia, particularmente como ferramenta de apoio à prospecção mineral e na elaboração de mapas geológicos mais precisos e detalhados.

**Palavras-chave:** *Machine Learning*; Corpos Máficos-Ultramáficos; Complexo Anápolis-Itauçu; Sequência Anicuns-Itaberaí; Algoritmo *Random Forest*.

## ABSTRACT

This research investigates the application of Machine Learning in identifying mafic-ultramafic bodies within a 200 km<sup>2</sup> area in the southern portion of the Brasília Belt, focusing specifically on the boundary between the Anicuns-Itaberaí sequence and the Anápolis-Itaçu Complex, near the municipality of Anicuns, Goiás.. The study applied the Random Forest algorithm to process and analyze geophysical and remote sensing data, aiming to enhance the geological mapping and detection of these bodies. The research integrated aerogeophysical data, including magnetometry and radiometry, as well as satellite data, in order to develop a robust geospatial dataset. The model was trained using a significant fraction of the dataset to provide a predictive map of geological occurrences. The obtained results demonstrated the high effectiveness of the model, with significant accuracy in discriminating the geological bodies of interest, which were validated by comparisons with field data and provided new insights on the distribution and characteristics of mafic-ultramafic intrusive bodies in the studied area. This study highlights the potential of using Machine Learning techniques in geology, particularly as a supporting tool for mineral prospecting, as well as providing more accurate and detailed geological maps.

**Keywords:** Machine Learning; Mafic-Ultramafic Intrusive Bodies; Anápolis-Itaçu Complex; Anicuns-Itaberaí Sequence; Random Forest algorithm.

## ÍNDICE DE FIGURAS

Figura 1. Localização da área de estudo na porção sudoeste do estado de Goiás, abrangendo os municípios de Itaberaí, Anicuns e Americano do Brasil. SRC: WGS 1984. ....	3
Figura 2. Mapa esquemático da Faixa Brasília, destacando a área de estudo (modificado de Dardenne, 2000). ....	5
Figura 3. Mapa inicial de litologias interpretadas na área de estudo por Lacerda Filho et al. (2021) mostrando as localizações dos corpos máficos-ultramáficos.....	7
Figura 4. Exemplo de cinco variáveis utilizadas neste estudo: SRTM, eU, eTh, K% e ASA..	16
Figura 5. Distribuição dos pontos usados para treinamento ao longo da área de pesquisa. As cores são simbólicas e variam de acordo com a litologia amostrada. ....	26
Figura 6. Mapa preditivo gerado pelo modelo Random Forest indicando a classe de Corpos Máficos-Ultramáficos na área de estudo. ....	33
Figura 7. Mapa preditivo gerado pelo modelo Random Forest, indicando a distribuição predita de corpos máficos-ultramáficos na área de estudo. As subdivisões regionais para a apresentação dos resultados estão destacadas em polígonos numerados. Região 1: Corpo Córrego Seco; Região 2: Americano do Brasil; Região 3: Gongomé-Retiro; Região 4: Águas Claras de Goiás, Capelinha dos Gomes e Fazenda Paraíso; Regiões 5 e 6: Predições isoladas. ....	34
Figura 8. Mapa preditivo com enfoque na região do Corpo Córrego Seco (NP3δasc). O mapa foi gerado pelo modelo Random Forest e indica a probabilidade de ocorrência e os pontos de predição da classe de Corpos Máficos-Ultramáficos na área de estudo.....	35
Figura 9. (A) Afloramento de rochas gabróicas no leito do Rio dos Bois, com padrão de fraturamento em “X” nas direções NNE-SSW e ENE-WNW (Ponto: TFLB_20). (B) Gabro cortado por dique de diabásio com direção NW-SE. (Ponto TFLB_20). (C) Blocos de gabro de granulação média (TFLB_22). ....	37

Figura 10. (A) Fotografia aérea tirada através de drone. Rocha de composição granodiorítica disposta em lajedo. (Ponto: TFLB_19). (B) Mistura física entre fenocristais de feldspato de rocha félsica em rocha máfica. (Ponto TFLB_19). .....	38
Figura 11. Mapa preditivo com enfoque na região do Complexo Americano do Brasil (NP3 $\mu$ aa). O mapa foi gerado pelo modelo Random Forest e indica a probabilidade de ocorrência e os pontos de predição da classe de Corpos Máficos-Ultramáficos na área de estudo. ....	39
Figura 12. (A) Afloramento de blocos e matacões de composição ultramáficas (Ponto: TFLB_05). (B) Bloco de piroxenito exibindo textura inequigranular porfirítica e grau moderado de intemperismo (Ponto: TFLB_05). (C) Amostra de piroxenito (Ponto: TFLB_05). ....	41
Figura 13. Mapa preditivo com enfoque na região do Corpo Gongomé-Retiro (NP3 $\mu$ tr). O mapa foi gerado pelo modelo Random Forest e indica a probabilidade de ocorrência e os pontos de predição da classe de Corpos Máficos-Ultramáficos na área de estudo.....	42
Figura 14. (A) Afloramento de blocos e matacões de composição gabroica/diorítica(Ponto: TFLB_04). (B) Afloramento de rochas ultramáficas na vizinhança sudoeste do Corpo Gongomé-Retiro (Ponto: TFLB_05). ....	43
Figura 15. (A) Blocos e matacões de composição gabróica nas intermediações do Corpo Gongomé-Retiro (Ponto: TFLB_08). (B) Detalhe da imagem “(A)” exibindo os cristais de granada em uma das superfícies intemperizadas, formando um aspecto de “catapora”. ....	44
Figura 16. Mapa preditivo com enfoque na região dos corpos Águas Claras de Goiás, Fazenda Paraíso e Capelinha dos Gomes. O mapa foi gerado pelo modelo Random Forest e indica a probabilidade de ocorrência e os pontos de predição da classe de Corpos Máficos-Ultramáficos na área de estudo.....	46
Figura 17. Afloramento de rochas metabásicas do ponto TFLB_24.....	47
Figura 18. (A) Matacão de composição gabroica na porção centro-norte do Corpo Capelinha dos Gomes (Ponto: TFLB_09). (B) Detalhe da parte superior do matacão da imagem “(A)”, exibindo foliação formada pela orientação mineral (135/80), possível feição de borda.....	48

Figura 19. (A) Enclave máfico em rochas intermediárias com foliação subverticalizada (123/81) (Ponto TFLB_10). (B) Detalhe para a borda do enclave evidenciando a segregação de bandas melanocráticas e leucocráticas.....	49
Figura 20. Encosta de morro composta por blocos desarticulados de gabronoritos nas intermediações do Corpo Fazenda Paraíso (Ponto TFLB_11). .....	50
Figura 21. (A) Afloramento de metabasitos nas intermediações do Corpo Águas Claras de Goiás (Ponto: TFLB_13). (B) Metabasito (Ponto: TFLB_13). .....	51
Figura 22. (A) Micaxistos em processo de gnaissificação (Ponto TFLB_16). (B) Blocos de metadiorito (Ponto TFLB_16). .....	52
Figura 23. (A) Visão ampla do afloramento do ponto TFLB_14. (B) Blocos rolados de metagabronorito (TFLB_14). (C) Detalhe para o metagabronorito do ponto TFLB_14 com foliação evidente. ....	53
Figura 24. Ortognaisse básico milonitizado (Ponto TFLB_17) com detalhe para a foliação milonítica da amostra coletada no ponto. ....	54
Figura 25. (A) Blocos desarticulados de ortognaisse migmatizado (Ponto TFLB_17). (B) Detalhe para uma banda melanocrática dos ortognaissees do Ponto TFLB_17. ....	55
Figura 26. Afloramento de hornblenda-gabros do ponto TFLB_26. ....	56
Figura 27. Mapa preditivo com enfoque nas predições isoladas na região sudeste da área de estudo, indicando a probabilidade de ocorrência da classe de Corpos Máficos-Ultramáficos na área de estudo. ....	57
Figura 28. (A) Afloramento decamétrico composto por ortognaissees (Ponto TFLB_12). (B e C) Detalhe para os enclaves máficos contidos nos ortognaissees do ponto TFLB_12. ....	58
Figura 29. Mapa preditivo com enfoque nas predições isoladas no centro da área de estudo, destacando a identificação de corpos máficos-ultramáficos pelo modelo Random Forest. ....	59
Figura 30. (A) Blocos de hornblenda-gabronorito desarticulados (Ponto TFLB_15). ....	60

## ÍNDICE DE TABELAS

Tabela 1: Reclassificação das litologias para aplicação do algoritmo de <i>ML</i> . As litologias reclassificadas estão marcadas com um asterisco (*). .....	<b>18</b>
Tabela 2: Variáveis utilizadas para treinamento do modelo, com indicação de tipo (variáveis originais, variáveis derivadas de outras variáveis). .....	<b>19</b>
Tabela 3: Modelo geral de uma matriz de confusão. ....	<b>20</b>
Tabela 4: Variáveis utilizadas para treinamento do modelo, com indicação de tipo (variáveis originais, variáveis derivadas de outras variáveis).....	<b>25</b>
Tabela 5: Métricas de desempenho do algoritmo <i>Random Forest</i> .....	<b>28</b>
Tabela 6: <i>Ranking</i> de importância das variáveis utilizadas.....	<b>29</b>
Tabela 7: Matriz de Confusão exibindo a relação percentual entre as litologias reais e previstas pelo modelo. A classe de Corpos M-UM está evidenciada por um contorno preto.....	<b>31</b>

## LISTA SIGLAS E ABREVIATURAS

ASA – Amplitude do Sinal Análítico

AML – Algoritmo de *Machine Learning*

AUC – *Area Under the Curve*

M-UM – Máfico-Ultramáfico

Ma – Milhões de anos

ML – *Machine Learning*

SHP – *Shapefile*

SRTM – *Shuttle Radar Topography Mission*

SVM – *Support Vector Machines*

TIFF – *Tagged Image File Format*

USGS – *United States Geological Survey*

MCC – *Matthews Correlation Coefficient*

ROC – *Receiver Operating Characteristic*

SRC – Sistema de Referência de Coordenadas

## SIGLAS DE UNIDADES GEOLÓGICAS

- SAI - Sequência Metavulcanosedimentar Anicuns-Itaberaí
- CAI – Complexo Granulítico Anápolis Itauçu
- CAB – Complexo Máfico-Ultramáfico Americano do Brasil
- N1dl – Coberturas detrito-lateríticas ferruginosas
- NP1 $\gamma$ cl – Granito Córrego da Lavrinha
- NP1 $\gamma$ ms – Metagranito Sanclerlândia
- NP1 $\gamma$ sj – São João
- NP1aibu – Sequência Anicuns-Itaberaí, básico- ultrabásica
- NP1aibutx – Sequência Anicuns-Itaberaí, talco xisto
- NP1aipp – Sequência Anicuns-Itaberaí, psamo-pelítica
- NP1aippmx – Sequência Anicuns-Itaberaí, mica-xisto
- NP1aippqt – Sequência Anicuns-Itaberaí, quartzito
- NP1airch – Sequência Anicuns-Itaberaí, metachert
- NP1aircmm – Sequência Anicuns-Itaberaí, mármore
- NP1aircr – Sequência Anicuns-Itaberaí, metarritmito
- NP1aito – Complexo Granulítico Anápolis-Itauçu, ortogranulito
- NP1aitoa – Complexo Granulítico Anápolis-Itauçu, granulito ácido
- NP1aitob – Complexo Granulítico Anápolis-Itauçu, granulito básico
- NP1aitp – Complexo Granulítico Anápolis-Itauçu, paraganulito
- NP2 $\gamma$ c – Granito Choupana
- NP2cam – Metagranito Campestre
- NP2camg – Metagranito Campestre
- NP23 $\gamma$ t – Metagranito Turvânia
- NP $\gamma$ l – Granito Lage
- NP $\gamma$ sr – Metadiorito Córrego Santa Rosa
- PP4MP1mos – Mossâmedes, sedimentar
- NP3 $\mu$ aa – Corpo Americano do Brasil
- NP3 $\mu$ ta – Águas Claras de Goiás

NP3 $\mu$ tc – Capelinha dos Gomes

NP3 $\mu$ tf – Fazenda Paraíso

NP3 $\mu$ tr – Gongomé-Retiro

NP3 $\delta$ asc $\delta$ i – Córrego Seco, metadiorito

NP3 $\delta$ ascg – Córrego Seco, metagabro

## SUMÁRIO

1. INTRODUÇÃO.....	1
1.1. HISTÓRICO E JUSTIFICATIVA .....	1
1.2. OBJETIVOS.....	1
2. LOCALIZAÇÃO DA ÁREA DE ESTUDO E VIAS DE ACESSO .....	2
2.1 LOCALIZAÇÃO.....	2
2.2 ACESSO.....	2
3. REVISÃO DE LITERATURA .....	4
3.1 CONTEXTO GEOLÓGICO REGIONAL.....	4
3.1.1 PROVÍNCIA TOCANTINS.....	4
3.1.2 FAIXA BRASÍLIA .....	4
3.1.3 ARCO MAGMÁTICO DE GOIÁS (AMG) .....	5
3.1.4 ARCO MAGMÁTICO DE ARENÓPOLIS.....	6
3.1.5 CORPOS INTRUSIVOS MÁFICOS-ULTRAMÁFICOS.....	7
3.1.5.1 CORPO CÓRREGO SECO .....	8
3.1.5.2 CORPO AMERICANO DO BRASIL.....	8
3.1.5.3 CORPO ÁGUAS CLARAS DE GOIÁS.....	9
3.1.5.4 CORPO CAPELINHA DOS GOMES .....	9
3.1.5.5 CORPO GONGOMÉ-RETIRO.....	9
3.1.5.6 CORPO FAZENDA PARAÍSO .....	10
3.2 MACHINE LEARNING NA GEOLOGIA.....	10
3.3 ALGORITMOS DE APRENDIZADO DE MÁQUINA.....	10
3.3.1 REDES NEURAIIS ARTIFICIAIS .....	10
3.3.2 MÁQUINAS DE VETORES DE SUPORTE (SVM).....	11
3.3.3 ÁRVORES DE DECISÃO .....	11
3.3.4 ALGORITMOS DE ENSEMBLE: <i>RANDOM FOREST</i> .....	11
3.4 RANDOM FOREST .....	11
3.5 ESTUDOS DE CASO RELEVANTES .....	12
4. METODOLOGIA.....	13
4.1 AQUISIÇÃO DE DADOS INDIRETOS .....	13

4.1.1 DADOS GEOLÓGICOS .....	13
4.1.2 IMAGENS DE SATÉLITE .....	13
4.1.3 MODELO DIGITAL DE ELEVAÇÃO (MDE) .....	14
4.1.4 DADOS AEROGEOFÍSICOS .....	14
4.2 AQUISIÇÃO DE DADOS POR CAMPANHA DE CAMPO .....	15
4.3 TRATAMENTO DOS DADOS .....	15
4.3.1 NORMALIZAÇÃO .....	16
4.3.2 IMPUTAÇÃO DE VALORES AUSENTES .....	17
4.3.3 RECLASSIFICAÇÃO DAS LITOLOGIAS .....	17
4.4 ENGENHARIA DE RECURSOS .....	18
4.5 MODELAGEM .....	19
4.5.1 ESCOLHA DO ALGORITMO .....	19
4.5.2 VALIDAÇÃO CRUZADA E OTIMIZAÇÃO DE HIPERPARÂMETROS .....	19
4.6 AVALIAÇÃO DO MODELO .....	20
4.6.1 MÉTRICAS DE DESEMPENHO DO MODELO .....	20
4.6.2 MÉTRICAS DE IMPORTÂNCIA DAS VARIÁVEIS .....	22
4.6.2.1 GANHO DE INFORMAÇÃO ( <i>INFOGAIN</i> ) .....	22
4.6.2.2 TAXA DE GANHO ( <i>GainRatio</i> ) .....	23
4.6.2.3 ÍNDICE GINI .....	23
4.6.2.4 QUI-QUADRADO ( $\chi^2$ ) .....	24
4.7 INTEGRAÇÃO E AMOSTRAGEM DE DADOS .....	24
4.8 TREINAMENTO E APLICAÇÃO DO MODELO <i>RANDOM FOREST</i> .....	27
4.9 CLASSIFICAÇÃO DAS AMOSTRAS DE CAMPO .....	27
5. RESULTADOS E DISCUSSÕES .....	28
5.1 AVALIAÇÃO DO MODELO .....	28
5.1.1 MÉTRICAS DE DESEMPENHO .....	28
5.1.2 IMPORTÂNCIA DAS VARIÁVEIS .....	29
5.1.3 MATRIZ DE CONFUSÃO .....	31
5.2 PREDIÇÕES DO MODELO .....	32
5.2.1 MAPA PREDITIVO PARA CORPOS MÁFICOS-ULTRAMÁFICOS .....	32
5.3 VALIDAÇÃO DOS DADOS EM CAMPO .....	33

5.3.1 REGIÃO 1: SW DA ÁREA DE ESTUDO - CORPO CÓRREGO SECO .....	34
5.3.2 REGIÃO 2: NW DA ÁREA DE ESTUDO - CORPO AMERICANO DO BRASIL.....	38
5.3.3 REGIÃO 3: N DA ÁREA DE ESTUDO - CORPO GONGOMÉ-RETIRO .....	41
5.3.4 REGIÃO 4: E DA ÁREA DE ESTUDO - CORPOS ÁGUAS CLARAS DE GOIÁS, CAPELINHA DOS GOMES E FAZENDA PARAÍSO .....	45
5.3.5 REGIÃO 5: SE DA ÁREA DE ESTUDO - PREDIÇÕES ISOLADAS.....	56
5.3.6 REGIÃO 6: CENTRO DA ÁREA DE ESTUDO - PREDIÇÕES ISOLADAS .....	58
6. CONSIDERAÇÕES FINAIS .....	60
7. REFERÊNCIAS BIBLIOGRÁFICAS .....	62
7.1 REFERÊNCIAS DE <i>SOFTWARES</i> UTILIZADOS .....	65
REFERÊNCIAS DE PLATAFORMAS UTILIZADOS .....	65
8. APÊNDICES .....	66
APÊNDICE A: MATRIZ DE CONFUSÃO EXPANDIDA.....	66
APÊNDICE B: TABELA DE PONTOS .....	67
APÊNDICE C: MAPA DE PONTOS DA ÁREA DE PESQUISA 1:100000 .....	71
APÊNDICE D: MAPA PREDITIVO DA ÁREA DE PESQUISA 1:100000 .....	72

# 1. INTRODUÇÃO

## 1.1. HISTÓRICO E JUSTIFICATIVA

A região do contato entre o Arco Magmático Arenópolis e o Complexo Anápolis-Itauçu, localizada na porção sul do Estado de Goiás, é uma área de significativa complexidade geológica. Esta região é marcada pela interação entre um complexo granulítico e uma sequência supracrustal, com a presença marcante de corpos intrusivos de tonalitos, granodioritos, granitos e rochas máficas-ultramáficas. Estes corpos máficos-ultramáficos são cruciais para a compreensão da evolução geológica e metalogenética da área, representando também um potencial significativo para exploração mineral (Almeida, 1967; Fuck *et al.*, 1994). A identificação precisa e a caracterização desses corpos rochosos são desafiadoras devido à complexidade geológica da região e à sobreposição de eventos tectônicos e metamórficos (Pimentel *et al.*, 2004; Dardenne, 2000).

Nesse contexto, as técnicas de *Machine Learning* emergem como ferramentas promissoras, capazes de processar e analisar grandes volumes de dados geofísicos e geológicos de forma mais eficiente e precisa do que os métodos tradicionais (Tabosa, 2023; Costa *et al.*, 2019, 2020; Kuhn *et al.*, 2019). O uso do algoritmo *Random Forest*, em particular, oferece uma abordagem robusta para a identificação e classificação de litologias, permitindo uma melhor discriminação dos corpos máficos-ultramáficos em meio à complexidade regional (Goodfellow *et al.*, 2015).

## 1.2. OBJETIVOS

O objetivo principal deste trabalho foi empregar técnicas de *Machine Learning*, especificamente o algoritmo *Random Forest*, para refinar a caracterização de rochas de composição semelhante à dos corpos máficos-ultramáficos. O foco foi aprimorar a compreensão desses corpos já mapeados, melhorando nossa habilidade de detalhar suas extensões, composições e inter-relacionamentos geológicos dentro do contexto do Arco Magmático Arenópolis e do Complexo Anápolis-Itauçu. Para isso, integramos e analisamos dados aerogeofísicos, litológicos e de sensoriamento remoto, visando uma abordagem mais precisa e detalhada no mapeamento geológico.

Espera-se que o estudo contribua para determinar a localização, características geométricas e possíveis novas ocorrências desses corpos intrusivos.

Como objetivos específicos pretendeu-se:

- **Coletar e processar dados geofísicos, topográficos e geológicos** relevantes para a área de estudo.
- **Preparar um conjunto de dados abrangente** que integre variáveis geofísicas e geológicas essenciais para a análise.
- **Treinar e validar o modelo de *Machine Learning*** para prever a presença e características dos corpos máficos-ultramáficos.
- **Avaliar a importância das variáveis** no modelo para entender os fatores críticos que influenciam a identificação desses corpos.
- **Gerar um mapa preditivo** da região com foco nos corpos identificados, facilitando a visualização e a interpretação dos dados.
- **Comparar os resultados obtidos** com observações de campo para validar a eficácia do modelo e discutir suas implicações geológicas e práticas.

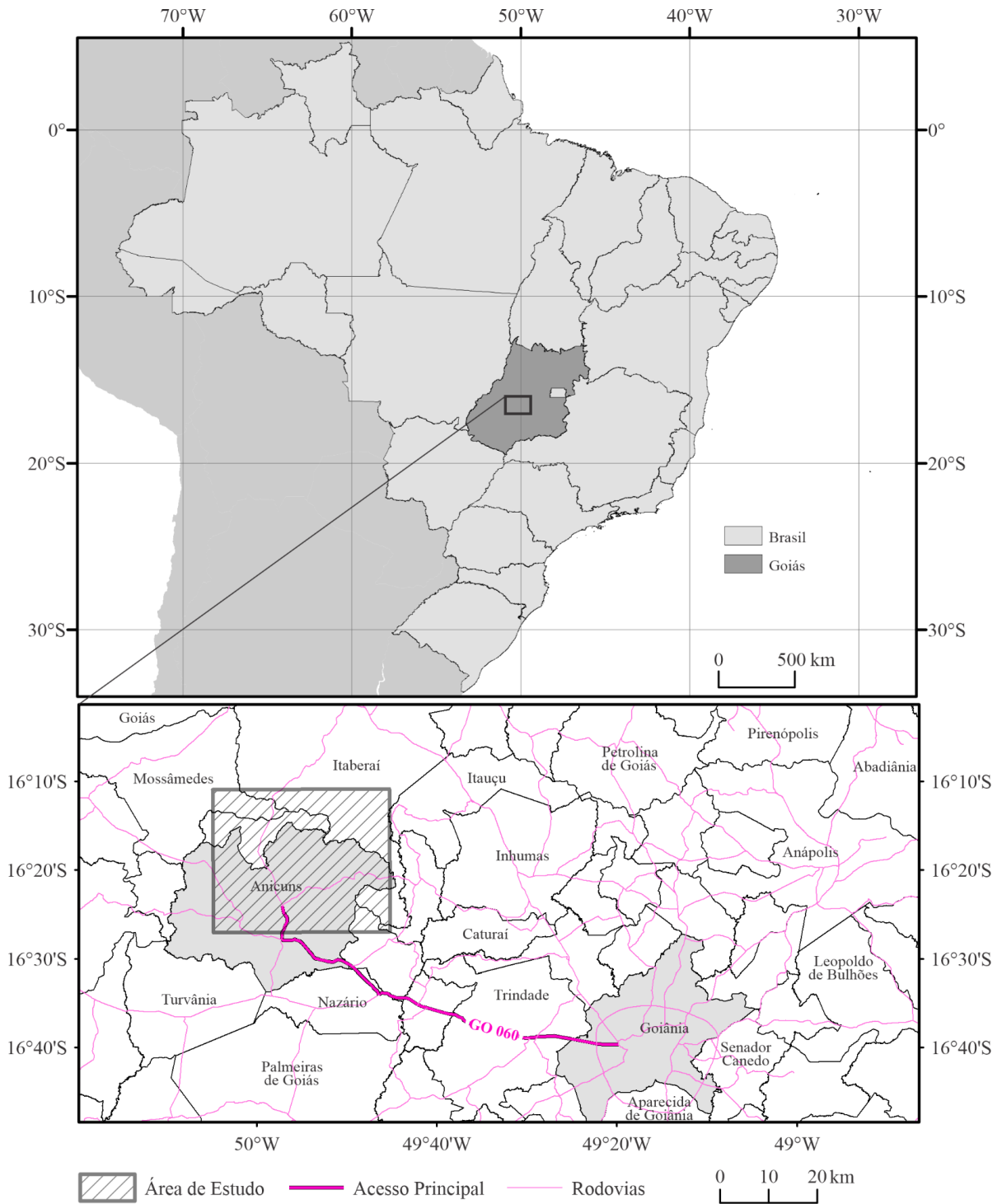
## **2. LOCALIZAÇÃO DA ÁREA DE ESTUDO E VIAS DE ACESSO**

### **2.1 LOCALIZAÇÃO**

A área de estudo possui aproximadamente 200 km<sup>2</sup> e está localizada na porção centro-oeste do Estado de Goiás, englobando os municípios de Itaberaí, Anicuns e Americano do Brasil (Figura 1). Geologicamente, está situada na porção sul do Arco Magmático de Goiás, inserido no contexto geológico da Faixa Brasília.

### **2.2 ACESSO**

O acesso à área de estudo pode ser realizado a partir da capital do Estado, Goiânia, seguindo em sentido oeste pela rodovia GO-060 (Figura 1). A região é acessível por estradas pavimentadas e vicinais, facilitando a logística para as campanhas de campo. As principais vias de acesso permitem um deslocamento eficiente, garantindo a viabilidade das atividades de coleta de dados e amostras no campo.



**Figura 1. Localização da área de estudo na porção sudoeste do estado de Goiás, abrangendo os municípios de Itaberaí, Anicuns e Americana do Brasil. SRC: WGS 1984. Fonte: IBGE, 2024.**

Durante a estação chuvosa, algumas estradas vicinais podem apresentar condições adversas, tornando o acesso mais difícil. Recomenda-se o uso de veículos 4x4 para garantir a mobilidade em todas as épocas do ano.

### **3. REVISÃO DE LITERATURA**

#### **3.1 CONTEXTO GEOLÓGICO REGIONAL**

##### **3.1.1 PROVÍNCIA TOCANTINS**

A Província Tocantins é um complexo sistema orogênicos brasileiros gerado principalmente pela colisão dos crátons São Francisco, Amazonas e Paranapanema (Almeida, 1967). Essa província é caracterizada por três cinturões de dobramentos e empurrões denominados Faixas Brasília, Paraguai e Araguaia.

##### **3.1.2 FAIXA BRASÍLIA**

A Faixa Brasília (FB) (Figura 2) estende-se por cerca de 1.000 km de extensão, por toda a margem oeste do Cráton do São Francisco. Representa a parte oriental/central da Província Tocantins (Almeida, 1967). A FB é reconhecida como um orógeno acrescionário complexo, composto por uma série de unidades tectônicas que se desenvolveram a partir de processos colisionais.

Este orógeno é subdividido em cinco porções distintas: (1) a zona de antepaís, (2) a zona externa, (3) a zona interna, (4) o Maciço de Goiás e (5) o Arco Magmático de Goiás. A intensidade do metamorfismo na Faixa Brasília aumenta de leste a oeste (Figura 2), desde rochas com baixo grau metamórfico próximas ao Cráton São Francisco até rochas de alto gradiente metamórfico na região central. A evolução das deformações sugere uma vergência tectônica rumo ao Cráton São Francisco (Dardenne, 2000; Fuck *et al.*, 1994). A área de estudo está inserida no contexto do Arco Magmático de Goiás.

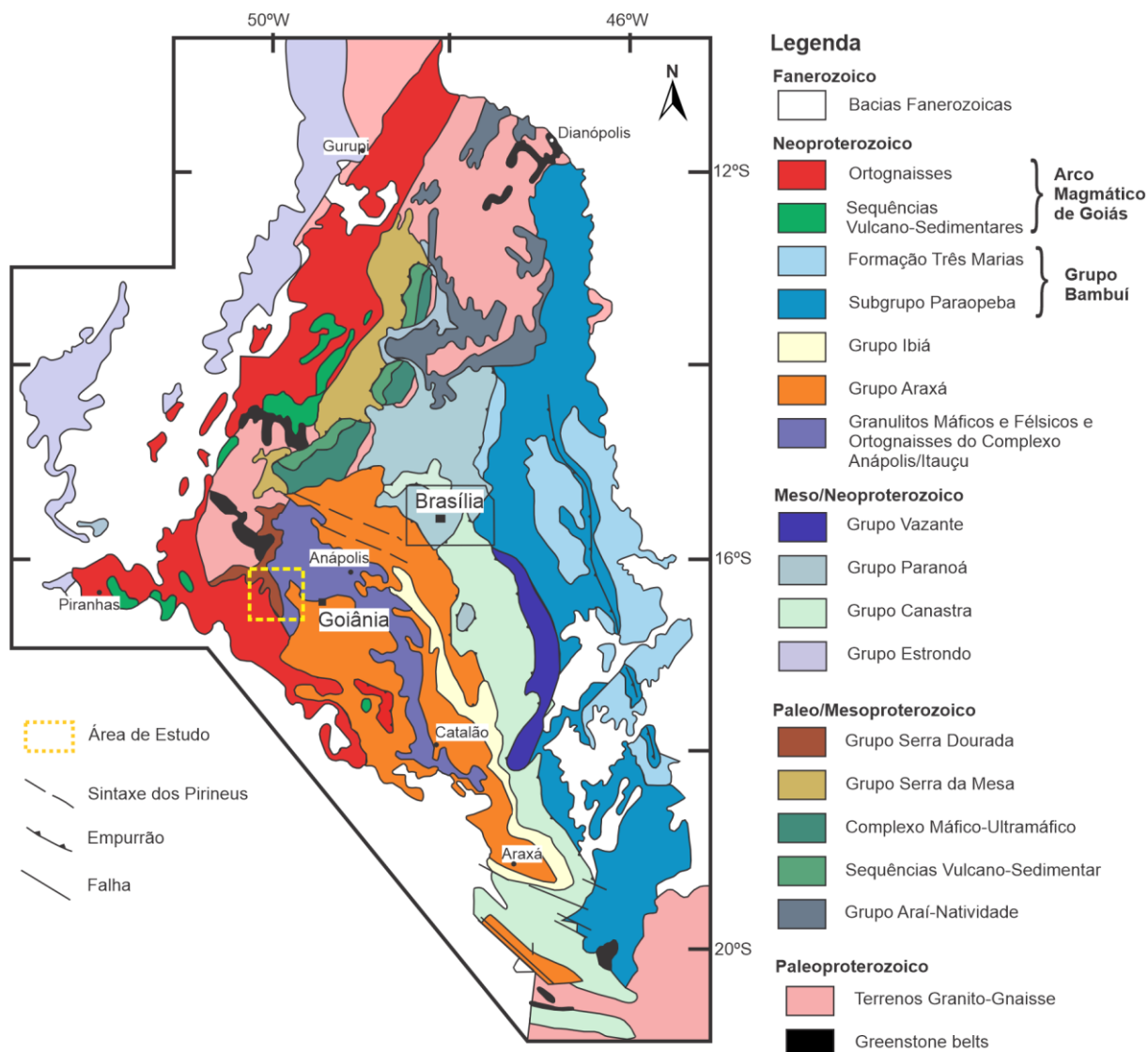


Figura 2. Mapa esquemático da Faixa Brasília, destacando a área de estudo (modificado de Dardenne, 2000).

### 3.1.3 ARCO MAGMÁTICO DE GOIÁS (AMG)

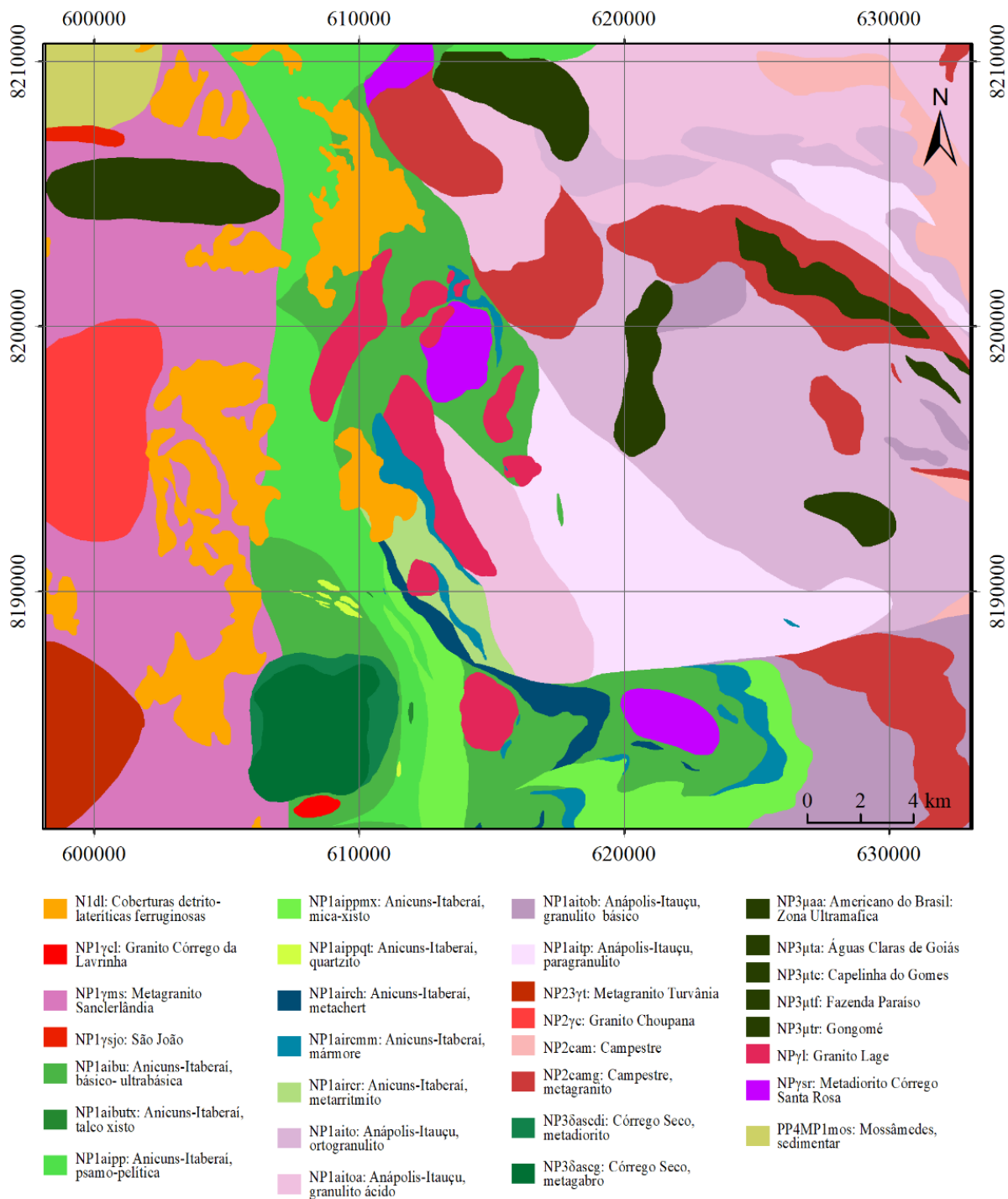
O Arco Magmático de Goiás, situado a oeste da Faixa Brasília, é interpretado como um complexo de arcos vulcânicos intra-oceânicos formados durante a fase de subducção entre 890-860 Ma, posteriormente acrescidos à Faixa Brasília (Pimentel *et al.*, 2004). Este arco é composto por importantes unidades litoestratigráficas que incluem as sequências Arenópolis-Piranhas, Anicuns-Itaberaí e Adelândia (ca. 850 Ma), formadas durante a fase insular do arco magmático, e as sequências Bom Jardim de Goiás e Jaupaci (ca. 750 Ma), associadas ao

magmatismo cálcio-alcálico em ambientes de arco intra-oceânico e margem continental (Lacerda Filho *et al.*, 2021).

Composto principalmente por rochas cálcio-alcálicas de idade neoproterozoica, incluindo tonalitos, dioritos cálcio-alcálicos e rochas vulcânicas, o Arco Magmático de Goiás é dividido em dois segmentos: o Arco Magmático de Mara Rosa, a norte, e o Arco Magmático de Arenópolis, a sul (Viana *et al.*, 1995; Pimentel *et al.*, 1997).

### **3.1.4 ARCO MAGMÁTICO DE ARENÓPOLIS**

A área de estudo se localiza no domínio do Arco Magmático Arenópolis (AMA) (Fuck *et al.*, 1994) (Figura 3). O AMA (830 - 600 Ma) compreende seis sequências metavulcanossedimentares de diferentes idades, de oeste para leste: Bom Jardim de Goiás (ca. 750 Ma), Arenópolis-Piranhas (ca. 929 Ma), Iporá-Amorinópolis (ca. 630 Ma), Jaupaci (ca. 747 Ma), Bacilândia (ca. 590 Ma), Adelândia (ca. 820 Ma) e Anicuns-Itaberaí (ca. 830 Ma). São separadas entre si por rochas metaplutônicas de composição tonalítica a granodiorítica, formadas em ambiente de arcos insulares e continentais (Laux *et al.*, 2004, 2005; Pimentel e Fuck, 1992; Barbosa, 1987; Guimarães *et al.*, 2012a; Marques *et al.*, 2017; Rodrigues, Gioia e Pimentel, 1999; Navarro, 2006, 2007; apud Lacerda Filho, 2021). Seu setor oriental abrange rochas sin-colisionais ortoderivadas (800 - 634 Ma) e rochas máfico-ultramáficas tardi colisionais de ~630 Ma (Silva *et al.*, 2011). Associadas ao arco ocorrem intrusões pós-colisionais máficas e félsicas (620 - 500 Ma) (Silva *et al.*, 2011). A cristalização destas rochas ocorreu em dois períodos. O primeiro entre ca. 620 a 560 Ma abrangendo as rochas máfico-ultramáficas e poucas intrusões félsicas. O segundo ocorre por volta de ca. 510 Ma abrangendo rochas félsicas.



**Figura 3. Mapa inicial de litologias interpretadas na área de estudo por Lacerda Filho et al. (2021) mostrando as localizações dos corpos máficos-ultramáficos.**

### 3.1.5 CORPOS INTRUSIVOS MÁFICOS-ULTRAMÁFICOS

Dentre os principais corpos máficos-ultramáficos da região, destacam-se os corpos Córrego Seco, Americano do Brasil, Gongomé, Águas Claras, Capelinha dos Gomes e Fazenda Paraíso.

### **3.1.5.1 CORPO CÓRREGO SECO**

O Corpo Córrego Seco (NP3δasc) é identificado por sua geometria circular distinta, observada através de levantamentos aerogeofísicos. Este corpo intrusivo apresenta contatos intrusivos com as sequências supracrustais e apresenta uma diversidade litológica, que inclui hornblenda leucogabros, metagabros e noritos. As rochas do Corpo Córrego Seco frequentemente exibem evidências de alteração hidrotermal, com a formação de minerais secundários como magnetita e clorita (Lacerda Filho *et al.*, 2021).

O Corpo Córrego Seco foi cartografado por Araújo (1997, apud Lacerda Filho, 2021) como um *stock* semi-circular com cerca de 30 km<sup>2</sup>, localizado a norte de Anicuns. Araújo subdividiu este corpo em fácies gabro na porção central, constituído por metagabros, metanoritos e anortositos, sequenciado por metaquartzo dioritos, quartzo dioritos, metadioritos, microdioritos e anfíbolitos nas bordas. Os gabros foram descritos nas margens do Rio dos Bois na Fazenda Cachoeirinha, exibindo coloração cinza avermelhada, com granulação média e estruturalmente fraturada. O relevo do Corpo Córrego Seco é positivo, escarpado e rugoso, condicionado por drenagens espaçadas nas imagens SRTM (Araújo, 1997; apud Lacerda Filho, 2021).

### **3.1.5.2 CORPO AMERICANO DO BRASIL**

O Complexo Máfico-Ultramáfico Americano do Brasil (CAB) é uma intrusão máfica-ultramáfica acamadada que apresenta importantes processos magmáticos e potencial para mineralizações econômicas de Ni-Cu (Silva, 2009). O CAB é composto por uma sequência acamadada leste-oeste com aproximadamente 12 km de comprimento por 2 km de largura, formada por duas sequências distintas (Norte e Sul), interpretadas como duas câmaras magmáticas distintas, colocadas lado a lado ao longo de uma zona de falha. As sequências têm séries de cristalização similares, características de magmas toleíticos.

A sequência Norte é representada por intercalações de rochas ultramáficas cumuláticas (dunito, wehrlito, lherzolito e websterito) e gabronoritos intercalados. A sequência Sul é composta por gabronoritos intercalados com ultramáficas cumuláticas, principalmente websterito e lherzolito. As duas sequências são separadas pela falha normal Salgado, que possui direção leste-oeste. Falhas de cisalhamento de direção noroeste-sudeste deslocam as camadas e os corpos

mineralizados por centenas de metros (Nilson, 1981b apud Lacerda Filho 2021; Laux *et al.*, 2004; Pimentel *et al.*, 2004; Laux, 2004).

### **3.1.5.3 CORPO ÁGUAS CLARAS DE GOIÁS**

Os corpos Águas Claras foram inicialmente referenciados por Bressan (1977, apud Lacerda Filho *et al.*, 2021) e denominados em decorrência de sua área-tipo, como Córrego Águas Claras. Foram cartografados três *stocks* com comprimento máximo de 8,5 km, formando uma faixa noroeste na porção centro-oeste da Folha Itaberaí. Estes corpos estabelecem contatos intrusivos em metagranitos do Complexo Gnáissico Migmatítico Campestre e ortogranulitos do Complexo Anápolis-Itauçu.

Os principais afloramentos são representados por gabros, peridotitos serpentinizados, piroxenitos, anfíbolitos e talco-clorita xistos. Os gabros ocorrem intercalados com peridotitos serpentinizados e piroxenitos. Talco-clorita xistos exibem coloração amarela, vermelha e esbranquiçada (Nilson, 1992 apud Lacerda Filho *et al.*, 2021).

### **3.1.5.4 CORPO CAPELINHA DOS GOMES**

Segundo Marini *et al.* (1984a, apud Lacerda Filho, 2021), os terrenos granulitizados do oeste de Goiás, que englobam o Corpo Capelinha dos Gomes, são constituídos por supracrustais e ortogranulitos, gnáissico-migmatíticos e rochas intrusivas granitoides, cortados por *stocks* básico-ultrabásicos ediacaranos. O Corpo Capelinha dos Gomes, associado à Suíte Taquaral-Trindade, distribui-se na parte central da Folha Itaberaí, com cerca de 9 km de comprimento, localizado no distrito de Santa Maria. Subdividido em litofácies metapiroxenito e metagabro, inclui peridotitos serpentinizados, metagabros intercalados com piroxenitos, metanoritos, anfíbolitos, piroxenitos e talco-clorita xistos (Araújo, 1997; Winge, 1995a; apud Lacerda Filho 2021).

### **3.1.5.5 CORPO GONGOMÉ-RETIRO**

O Corpo Gongomé, delimitado por Pfrimer, Cândia e Teixeira (1981, apud Lacerda Filho, 2021), está localizado nas proximidades de Congomé e São José do Retiro, na Serra de Gongomé. Representado por um *stock* lenticular com 6,2 km de comprimento, é constituído por núcleo de metagabro uralitizado e biotitizado, com fenocristais de plagioclásio e granada,

transicionando para bordas de diorito, quartzo diorito, muscovita-zoizita-granada anfibolito e granada anfibolito fino. Dados Rb-Sr em gabrodiorito indicam idade de cristalização de  $615 \pm 108$  Ma (Winge, 1995b; Araújo, 1997; apud Lacerda Filho 2021).

### **3.1.5.6 CORPO FAZENDA PARAÍSO**

O Corpo Fazenda Paraíso foi cartografado como um *stock* com 3,6 km de comprimento, localizado 2,8 km a leste da localidade Capelinha, Município de Itaberaí-GO, e exibe relação de contato do tipo intrusivo nos ortogranulitos do Complexo Anápolis-Itaçu. Litotipos incluem metagabros, metanoritos e leucogabros, distribuídos na região dos Alves, trecho Capelinha-Araçu. O afloramento corresponde a uma rocha máfica, melanocrática a ultramelanocrática, com cristais de piroxênio e anfibólio, exibindo textura inequigranular de granulação média (Araújo, 1997; apud Lacerda Filho *et al.*, 2021).

## **3.2 MACHINE LEARNING NA GEOLOGIA**

A aplicação de técnicas de *Machine Learning* na geologia tem ganhado destaque nos últimos anos, devido à capacidade desses métodos de lidar com grandes volumes de dados e identificar padrões complexos (Kuhn *et al.* 2019, Costa *et al.*, 2019; Tabosa, 2023). O *Machine Learning* permite a análise integrada de diferentes tipos de dados geológicos, geofísicos e topográficos, facilitando a identificação de litologias, estruturas geológicas e potenciais depósitos minerais. Por exemplo, estudos de Rodriguez-Galiano *et al.* (2012) demonstraram como técnicas de aprendizado de máquina podem ser aplicadas para melhorar o mapeamento geológico em diversas regiões.

Diversas técnicas de *Machine Learning* têm sido aplicadas na geologia, incluindo redes neurais artificiais, máquinas de vetores de suporte (SVM), árvores de decisão e algoritmos de *ensemble*, como o *Random Forest*. Cada uma dessas técnicas possui características específicas que podem ser aproveitadas, dependendo do tipo de problema geológico em questão (Carranza & Laborte, 2015).

## **3.3 ALGORITMOS DE APRENDIZADO DE MÁQUINA**

### **3.3.1 REDES NEURAIAS ARTIFICIAIS**

As redes neurais artificiais são modelos computacionais inspirados na estrutura e funcionamento do cérebro humano. Elas são compostas por camadas de neurônios artificiais que processam informações e aprendem padrões a partir dos dados. Na geologia, as redes neurais têm sido utilizadas para a classificação de litologias, previsão de propriedades do solo e identificação de anomalias geofísicas (Goodfellow *et al.*, 2015; Bishop, 2006; Mitchell, 1997).

### **3.3.2 MÁQUINAS DE VETORES DE SUPORTE (SVM)**

As máquinas de vetores de suporte são algoritmos de classificação que buscam encontrar um hiperplano de separação ótimo entre diferentes classes de dados. (Lorena & de Carvalho, 2007).

### **3.3.3 ÁRVORES DE DECISÃO**

As árvores de decisão são modelos de classificação e regressão que utilizam uma estrutura hierárquica de decisões baseadas em valores dos atributos dos dados (Quinlan, 1987). Estas são amplamente utilizadas na geologia para a classificação de tipos de solo, previsão de deslizamentos de terra e identificação de áreas potenciais para mineração.

### **3.3.4 ALGORITMOS DE ENSEMBLE: *RANDOM FOREST***

O *Random Forest* é um algoritmo de *ensemble* que combina múltiplas árvores de decisão para melhorar a acurácia e robustez das previsões. Ele é particularmente eficaz na geologia devido à sua capacidade de lidar com grandes volumes de dados, tratar variáveis de diferentes tipos (numéricas e categóricas) e resistir ao *overfitting* (Breiman, 2001).

## **3.4 RANDOM FOREST**

O *Random Forest* é uma técnica de aprendizado de máquina que se destaca pela sua robustez e eficiência. Ele funciona criando diversas árvores de decisão a partir de diferentes subconjuntos de dados e atributos. Cada árvore de decisão é treinada em um subconjunto aleatório dos dados de treinamento e um subconjunto aleatório de atributos, selecionados com reposição (*bootstrap sample*). Este processo de "*bootstrap aggregating*" ou "*bagging*" garante a diversidade entre as árvores.

Após o treinamento, cada árvore de decisão gera uma predição. Para a classificação, a predição final do *Random Forest* é obtida através de uma votação majoritária entre todas as árvores, enquanto para a regressão, a média das predições individuais é utilizada. Este método é altamente eficiente na classificação de litologias e na identificação de padrões geológicos devido à sua capacidade de generalização e alta acurácia (Rodríguez-Galiano et al., 2012).

De acordo com Breiman (2001), o criador do Random Forest, as vantagens deste método incluem:

- **Robustez contra *Overfitting*:** A combinação de múltiplas árvores de decisão reduz significativamente o risco de *overfitting*. A diversidade introduzida pelo processo de *bagging* permite que o modelo generalize melhor em dados não vistos.
- **Capacidade de Lidar com Dados Heterogêneos:** O *Random Forest* pode lidar facilmente com variáveis contínuas e categóricas, tornando-o altamente versátil para diferentes tipos de problemas de modelagem.
- **Importância das Variáveis:** O *Random Forest* permite avaliar a importância relativa de cada variável no modelo. Esta avaliação é feita através do impacto que a alteração de uma variável específica tem na precisão do modelo, permitindo a identificação das variáveis mais influentes no processo de predição.

### 3.5 ESTUDOS DE CASO RELEVANTES

Diversos estudos têm demonstrado a eficácia do *Random Forest* em aplicações geológicas (Cracknell et al., 2014; Kuhn et al., 2018). Por exemplo, Tabosa (2023) utilizou o *Random Forest* para detectar corpos máficos-ultramáficos na região sul do Tocantins, demonstrando que o algoritmo pode identificar unidades geológicas mesmo em zonas de alta complexidade geológica. Costa et al. (2019) aplicaram técnicas de aprendizado de máquina para mapear litologias no Lineamento Cinzento, na Província Mineral de Carajás, Brasil, demonstrando a utilidade do *Random Forest* para a predição de litologias delimitadas por grandes estruturas geológicas, a partir de dados geofísicos.

Um estudo significativo realizado por Kuhn et al. (2019) no Cinturão de Cobre da África Central ilustra ainda mais a eficácia do *Random Forest* em mapeamentos litológicos. Os autores

realizaram uma série de estudos de caso que testaram a classificação litológica usando o algoritmo *Random Forest*, tanto em estágios iniciais de mapeamento quanto em refinamentos de mapas mais maduros. A pesquisa demonstrou que a combinação de técnicas de amostragem equilibrada e métricas de incerteza pode melhorar significativamente a precisão do mapeamento litológico. Este estudo destacou a capacidade do *Random Forest* em lidar com dados geológicos complexos e sublinhou a importância da entropia da informação, como uma métrica para avaliar a incerteza na classificação geológica.

## **4. METODOLOGIA**

### **4.1 AQUISIÇÃO DE DADOS INDIRETOS**

Os dados utilizados neste estudo foram obtidos de fontes públicas, incluindo o mapa geológico do Relatório Oeste de Goiás do Serviço Geológico do Brasil (SGB) (Lacerda Filho *et al.*, 2021), imagens de satélite do *Landsat 9*, modelo digital de elevação a partir da SRTM (*Shuttle Radar Topography Mission*) e produtos aerogeofísicos do Plano Aerogeofísico do Estado de Goiás. Esses dados foram integrados para fornecer uma visão abrangente das características geológicas da área de estudo.

#### **4.1.1 DADOS GEOLÓGICOS**

Os dados litológicos foram embasados no mapa geológico do oeste de Goiás (Lacerda Filho *et al.*, 2021), obtido através do Repositório Institucional do Serviço Geológico do Brasil (RiGeo). Este mapa inclui informações sobre as unidades litoestratigráficas, suas distribuições, contatos geológicos e relações estruturais. O arquivo *shapefile* com os polígonos das litologias foi recortado para o perímetro da área de estudo, que engloba 30 litologias no total.

#### **4.1.2 IMAGENS DE SATÉLITE**

Imagens de satélite do *Landsat 9* (especificamente as bandas 1 a 7), foram adquiridas através do *Earth Explorer* da USGS (*United States Geological Survey*), garantindo cobertura mínima de nuvens e correções prévias dos produtos. Estas imagens foram utilizadas para identificar diferentes tipos de cobertura do solo, feições estruturais e anomalias espectrais associadas a diferentes litologias.

#### 4.1.3 MODELO DIGITAL DE ELEVAÇÃO (MDE)

Para o modelo digital de elevação (MDE) foram utilizados dados de radar da missão SRTM (*Shuttle Radar Topography Mission*), obtidos através do *Earth Explorer* da *USGS*. Os dados de SRTM foram utilizados com o intuito de associar o relevo, declividade e aspectos do terreno ao modelo, úteis para uma análise geomorfológica durante a integração com outros dados geológicos/geofísicos.

#### 4.1.4 DADOS AEROGEOFÍSICOS

Os dados geofísicos, magnetométricos e gamaespectrométricos, foram processados por meio do *software Oasis Montaj* para eliminar ruídos e melhorar a qualidade das informações obtidas. Inicialmente armazenados em arquivos *geodatabase* (.gdb), os dados foram filtrados para um polígono que englobasse completamente a área de estudo, incluindo suas margens. Para evitar distorções, os pontos provenientes das linhas de controle, orientadas na direção Norte-Sul, foram removidos. Após a seleção dos dados, foram criados *grids* utilizando técnicas de filtragem para modelar a distribuição de propriedades físicas no subsolo. Especificamente, as técnicas de filtragem aplicadas foram:

- **ASA (Amplitude do Sinal Analítico):** Utilizou-se um filtro binomial para destacar as variações locais nos dados magnéticos, permitindo uma melhor identificação de anomalias geofísicas.
- **K% (Porcentagem de Potássio):** Aplicou-se uma interpolação por Curvatura Mínima para suavizar as variações nos dados radiométricos, garantindo uma representação mais fiel da distribuição de potássio no subsolo.
- **eTh (Tório equivalente) e eU (Urânio equivalente):** Ambos foram interpolados utilizando Curvatura Mínima para assegurar a continuidade espacial e a suavização dos dados, facilitando a interpretação dos resultados radiométricos.
- **Filtro 5x5:** Aplicado aos dados geofísicos para suavizar e eliminar ruídos residuais, melhorando a clareza e a precisão das informações.

Após a aplicação dessas técnicas de filtragem, os produtos resultantes foram utilizados para a amostragem descrita no item 4.7, permitindo uma análise detalhada das características geofísicas da área de estudo.

## **4.2 AQUISIÇÃO DE DADOS POR CAMPANHA DE CAMPO**

Além dos dados indiretos, foi realizada uma campanha de campo para coletar amostras em locais estratégicos dos corpos máficos/ultramáficos, os quais foram identificados com a ajuda dos resultados gerados pelo modelo preditivo. Utilizando-se um mapa com a delimitação dos corpos máficos-ultramáficos, a probabilidade de ocorrência desse grupo e os pontos preditos, foram definidos os sítios levando em consideração aglomerações de pontos de predição e a facilidade de acesso. Foram visitados pontos dentro e fora das delimitações destas litologias, utilizando-se uma escala de 1:100000.

A campanha incluiu a identificação e descrição de afloramentos chave, além da coleta de amostras de mão representativas das diferentes litologias presentes na área de estudo e medições de parâmetros geológicos (lineação mineral, acamamento, *strike* de camadas). Foram registradas informações iniciais das amostras e suas estruturas, relevo do local, tipo de afloramento e documentação fotográfica.

Os dados adquiridos em campo foram utilizados para validar as predições feitas pelo modelo de aprendizado de máquina. Esta validação envolveu a comparação das predições de litologias feitas pelo modelo com as observações de campo, como dimensões e origem do material (autóctone, alóctone) possibilitando verificar a precisão do modelo. As discrepâncias entre as predições do modelo e os dados de campo também foram analisadas para identificar suas possíveis causas e melhorar a interpretação geológica.

## **4.3 TRATAMENTO DOS DADOS**

Os dados coletados (Figura 4) passaram por um processo de tratamento para garantir sua qualidade e consistência. Este processo envolveu a normalização das variáveis, a imputação de valores ausentes e a reclassificação das litologias visando resultados mais significativos.

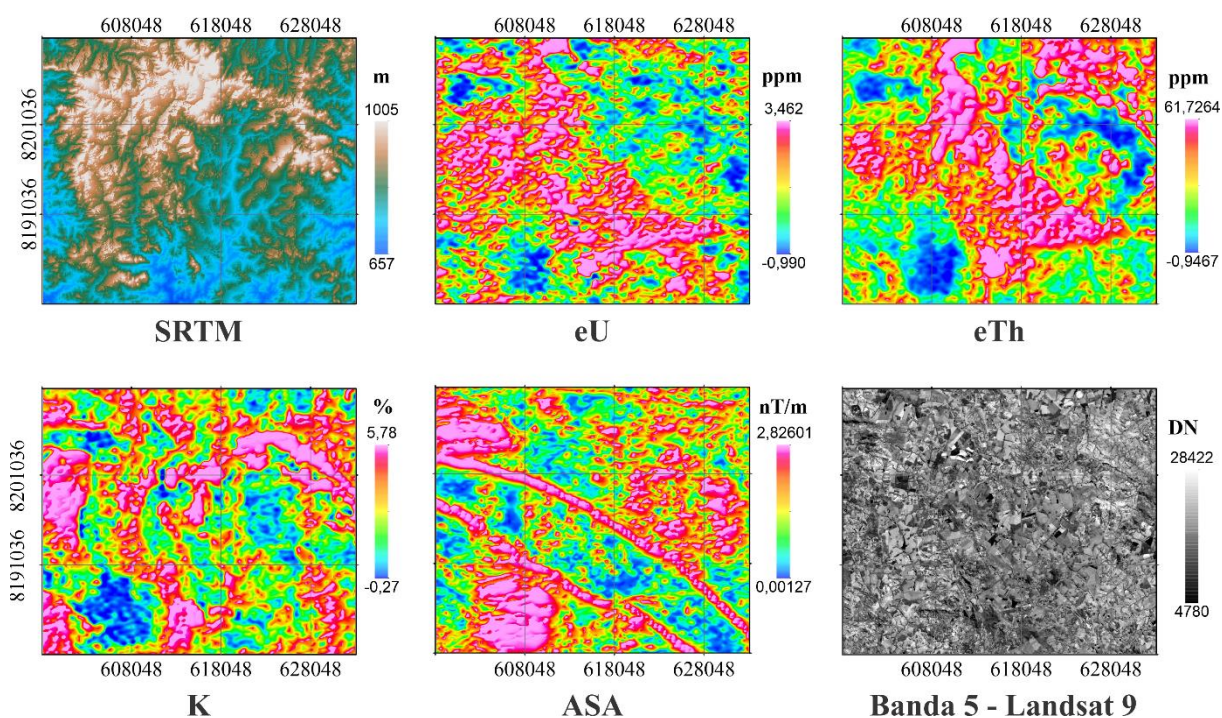


Figura 4. Exemplo de variáveis utilizadas neste estudo: SRTM, eU, eTh, K%, ASA e imagens de satélite.

### 4.3.1 NORMALIZAÇÃO

A normalização foi aplicada para ajustar os valores das variáveis em uma escala comum, facilitando a comparação e a análise multivariada. Para isso, foi utilizado o método de normalização *Z-score*, que ajusta os valores com base na média e no desvio padrão da variável:

$$Z = \frac{X - \mu}{\sigma}$$

Equação 1. Cálculo da normalização do tipo *Z-Score*. Fórmula utilizada para normalização das variáveis, ajustando os valores com base na média e no desvio padrão.

Onde:

- $Z$  é o valor normalizado.
- $X$  é o valor original da variável.
- $\mu$  é a média da variável.
- $\sigma$  é o desvio padrão da variável.

A escolha da normalização *Z-score* foi baseada na análise das características estatísticas das variáveis, que apresentaram grande dispersão e presença de *outliers*. A normalização *Z-score* é mais robusta contra *outliers* e preserva a distribuição original dos dados, proporcionando uma base mais estável para a aplicação do algoritmo *Random Forest*.

#### **4.3.2 IMPUTAÇÃO DE VALORES AUSENTES**

Os valores ausentes foram tratados utilizando a substituição destes pela mediana.

#### **4.3.3 RECLASSIFICAÇÃO DAS LITOLOGIAS**

Embora haja heterogeneidade nas fácies de cada litotipo, algumas litologias foram reclassificadas de modo a direcionar os resultados das predições para os corpos de composição máfica-ultramáfica. A reclassificação pode ser observada na Tabela 1.

**Tabela 1: Reclassificação das litologias para aplicação do algoritmo de *ML*. As litologias reclassificadas estão marcadas com um asterisco (\*). M-Mu: máfico-ultramáfico.**

<b>Sigla</b>	<b>Litologia/Unidade/Classe</b>	<b>Reclassificação</b>
* NP3 $\mu$ aa	Americano do Brasil - Zona Ultramáfica	Corpos M-UM
* NP3 $\mu$ ta	Águas Claras de Goiás	Corpos M-UM
* NP3 $\mu$ tc	Capelinha do Gomes	Corpos M-UM
* NP3 $\mu$ tf	Fazenda Paraíso	Corpos M-UM
* NP3 $\mu$ tr	Gongomé	Corpos M-UM
* NP3 $\delta$ asc $\delta$ i	Córrego Seco, metadiorito	Corpos M-UM
* NP3 $\delta$ asc $\delta$ g	Córrego Seco, metagabro	Corpos M-UM
N1dl	Coberturas detrito-lateríticas ferruginosas	N1dl
NP1aibu	Anicuns-Itaberaí, básico- ultrabásica	NP1aibu
NP1aibutx	Anicuns-Itaberaí, talco xisto	NP1aibutx
NP1aipp	Anicuns-Itaberaí, psamo-pelítica	NP1aipp
NP1aippmx	Anicuns-Itaberaí, mica-xisto	NP1aippmx
NP1aippqt	Anicuns-Itaberaí, quartzito	NP1aippqt
NP1airch	Anicuns-Itaberaí, metachert	NP1airch
NP1aircmm	Anicuns-Itaberaí, mármore	NP1aircmm
NP1aircr	Anicuns-Itaberaí, metarritmito	NP1aircr
NP1aito	Anápolis-Itauçu, ortogranulito	NP1aito
NP1aitoa	Anápolis-Itauçu, granulito ácido	NP1aitoa
NP1aitob	Anápolis-Itauçu, granulito básico	NP1aitob
NP1aitp	Anápolis-Itauçu, paragranulito	NP1aitp
NP1 $\gamma$ cl	Granito Córrego da Lavrinha	NP1 $\gamma$ cl
NP1 $\gamma$ ms	Metagranito Sanclerlândia	NP1 $\gamma$ ms
NP1 $\gamma$ sjó	São João	NP1 $\gamma$ sjó
NP23 $\gamma$ t	Metagranito Turvânia	NP23 $\gamma$ t
NP2cam	Campestre	NP2cam
NP2camg	Campestre, metagranito	NP2camg
NP2 $\gamma$ c	Granito Choupana	NP2 $\gamma$ c
NP $\gamma$ l	Granito Lage	NP $\gamma$ l
NP $\gamma$ sr	Metadiorito Córrego Santa Rosa	NP $\gamma$ sr
PP4MP1mos	Mossâmedes, sedimentar	PP4MP1mos

Os corpos intrusivos Americano do Brasil (NP3 $\mu$ aa), Águas Claras de Goiás (NP3 $\mu$ ta), Capelinha dos Gomes (NP3 $\mu$ tc), Fazenda Paraíso (NP3 $\mu$ tf), Gongomé (NP3 $\mu$ tr) e Córrego Seco (NP3 $\delta$ asc $\delta$ i e NP3 $\delta$ asc $\delta$ g) foram reclassificadas como corpos Máficos-Ultramáficos (M-UM), enquanto as demais litologias permaneceram com a sua sigla/classe original.

#### **4.4 ENGENHARIA DE RECURSOS**

A engenharia de recursos é uma etapa fundamental para melhorar o desempenho dos modelos de aprendizado de máquina. Nesta etapa, foram criadas variáveis derivadas a partir das variáveis geofísicas originais, como razões entre concentrações de elementos químicos, indicadores de anomalias geofísicas e a partir do MDE, como o índice de distância de drenagens. Essas variáveis são apresentadas na Tabela 2.

**Tabela 2: Variáveis utilizadas para treinamento do modelo, com indicação de tipo (variáveis originais, variáveis derivadas de outras variáveis).**

Variável	Un. de Medida	Resolução	Descrição	Tipo
K	%	100m	Porcentagem de Potássio	Original
eU	ppm	100m	Contagem de Urânio equivalente	Original
eTh	ppm	100m	Contagem de Tório equivalente	Original
SRTM	m	30m	Elevação	Original
ASA	nT/m	100m	Amplitude do Sinal Analítico	Original
B1	<i>Digital Number</i>	30m	Landsat 9 - Banda 1 - Azul Costeiro	Original
B2	<i>Digital Number</i>	30m	Landsat 9 - Banda 2 - Azul	Original
B3	<i>Digital Number</i>	30m	Landsat 9 - Banda 3 - Verde	Original
B4	<i>Digital Number</i>	30m	Landsat 9 - Banda 4 - Vermelho	Original
B5	<i>Digital Number</i>	30m	Landsat 9 - Banda 5 - Infravermelho Proximal	Original
B6	<i>Digital Number</i>	30m	Landsat 9 - Banda 6 - Infravermelho Médio	Original
B7	<i>Digital Number</i>	30m	Landsat 9 - Banda 7 - Infravermelho Distal	Original
IDD	-	10m	Índice de Distância de Drenagens	Derivada
eTh/K	Razão	100m	Proporção entre Tório e Potássio	Derivada
eU/K	Razão	100m	Proporção entre Urânio e Potássio	Derivada
ASA/eTh	Razão	100m	Proporção entre a ASA e Tório	Derivada
eU/eTh	Razão	100m	Proporção entre Urânio e Tório	Derivada
ASA/K	Razão	100m	Proporção entre a ASA e Potássio	Derivada
ASA/eU	Razão	100m	Proporção entre a ASA e Urânio	Derivada

## 4.5 MODELAGEM

A base teórica apresentada nos tópicos 4.5 e 4.6 tem como objetivo proporcionar uma compreensão aprofundada do funcionamento do algoritmo utilizado. É importante ressaltar que todos os cálculos e operações descritos são realizados pelo *software Orange Data Mining*, que oferece uma interface robusta e intuitiva para a análise de dados.

### 4.5.1 ESCOLHA DO ALGORITMO

O algoritmo *Random Forest* foi escolhido pelas suas características de robustez, eficiência e capacidade de interpretar a importância das variáveis. A escolha foi baseada na revisão da literatura e em considerações práticas sobre os dados disponíveis.

### 4.5.2 VALIDAÇÃO CRUZADA E OTIMIZAÇÃO DE HIPERPARÂMETROS

A validação cruzada é um método robusto para estimar a eficácia de modelos preditivos, especialmente útil para evitar o sobreajuste, garantindo que o modelo generalize bem para novos dados não vistos durante o treinamento (Hastie et al., 2009). Para otimizar o desempenho do modelo, foi utilizada validação cruzada e busca de hiperparâmetros (*Randomized Search*).

A validação cruzada foi realizada utilizando o método *k-fold*, com  $k = 3$ , para garantir que o modelo fosse treinado e testado em diferentes subconjuntos dos dados, minimizando o risco de *overfitting*. Desse modo, o conjunto foi separado em 3 partes, sendo 66% para treino e 33% para teste.

A otimização de hiperparâmetros através de métodos como *Randomized Search* é fundamental para encontrar a configuração ideal que maximize o desempenho do modelo preditivo (Bergstra & Bengio, 2012). Os hiperparâmetros ajustados incluíram o número de árvores na floresta, a profundidade máxima das árvores e o número mínimo de amostras por folha. Foram realizadas 10 iterações para encontrar a melhor combinação de parâmetros ( $n\_iter = 10$ ).

## 4.6 AVALIAÇÃO DO MODELO

### 4.6.1 MÉTRICAS DE DESEMPENHO DO MODELO

O modelo foi avaliado utilizando várias métricas de desempenho, incluindo a matriz de confusão, precisão, *recall* e *F1-score*.

- **Matriz de Confusão:** A matriz de confusão é uma ferramenta poderosa para visualizar o desempenho de um algoritmo de classificação, onde cada coluna da matriz representa o número de predições de cada classe, enquanto cada linha representa as instâncias na classe real (Kohavi & Provost, 1998).

Tabela 3: Modelo geral de uma matriz de confusão.

		Previsão	
		Positivo	Negativo
Real	Positivo	<i>TP</i>	<i>FN</i>
	Negativo	<i>FP</i>	<i>TN</i>

Onde:

- *TP (True Positives)* é o número de verdadeiros positivos.
- *TN (True Negatives)* é o número de verdadeiros negativos.
- *FP (False Positives)* é o número de falsos positivos.
- *FN (False Negatives)* é o número de falsos negativos.

Métricas como acurácia, precisão, recall e F1-score são essenciais para avaliar a qualidade de modelos de classificação, oferecendo uma visão holística da performance em termos de eficiência e confiabilidade (Powers, 2011).

- **Acurácia:** A acurácia é uma métrica fundamental para avaliar o desempenho do modelo de classificação. É calculada como a razão entre o número de predições corretas e o número total de amostras. A fórmula utilizada para calcular a acurácia é:

$$\mathbf{Acurácia} = \frac{TP+TN}{TP+TN+FP+FN}$$

**Equação 2. Cálculo da acurácia.**

- **Precisão (*Precision*):** Essa métrica evidencia a taxa de acerto dentro de todas as previsões positivas.

$$\mathbf{Precisão} = \frac{TP}{TP + FP}$$

**Equação 3. Fórmula para calcular a precisão, que indica a taxa de acerto dentro de todas as previsões positivas.**

- **Recall:** Essa métrica evidencia a taxa de acerto dentro de todos os valores positivos reais.

$$\mathbf{Recall} = \frac{TP}{TP + FN}$$

**Equação 4. Fórmula para calcular o recall, que indica a taxa de acerto dentro de todos os valores positivos reais.**

- **F1-Score:** Essa métrica é uma média entre as métricas Precisão e Recall.

$$\mathbf{F1\ Score} = \frac{\mathbf{Precisão} * \mathbf{Recall}}{\mathbf{Precisão} + \mathbf{Recall}}$$

**Equação 5. Fórmula para calcular o F1-Score, que é uma média harmônica entre precisão e recall.**

- **MCC:** O Coeficiente de Correlação de Matthews (MCC) é uma métrica de desempenho usada em problemas de classificação binária que leva em consideração todos os elementos da matriz de confusão. Ele é especialmente útil quando há um desbalanceamento entre as

classes, proporcionando uma medida equilibrada da qualidade das previsões do modelo (Stoica & Babu, 2023).

O MCC é calculado pela fórmula:

$$\text{MCC} = \frac{(TP \cdot TN) - (FP \cdot FN)}{\sqrt{(TP + FP) \cdot (TP + FN) \cdot (TN + FP) \cdot (TN + FN)}}$$

**Equação 6. Fórmula para calcular o Coeficiente de Correlação de Matthews (MCC).**

- Esta métrica varia de -1 a 1, onde:
- 1 indica uma previsão perfeita,
- 0 indica que o modelo não tem poder preditivo melhor do que um modelo aleatório,
- -1 indica uma previsão inversa perfeita.

#### 4.6.2 MÉTRICAS DE IMPORTÂNCIA DAS VARIÁVEIS

As métricas de importância das variáveis são ferramentas cruciais para avaliar o impacto de cada variável na previsão do modelo. No contexto do algoritmo *Random Forest*, utilizamos várias métricas para determinar a importância relativa de cada variável. As principais métricas usadas incluem Ganho de Informação, Taxa de Ganho, Índice Gini e Qui-Quadrado (Quinlan, 1986; Breiman et al., 1984; Agresti, 2002).

Ao considerar essas métricas, pode-se construir modelos mais robustos e interpretáveis, garantindo que as variáveis mais significativas sejam utilizadas de forma eficiente para melhorar a precisão das previsões. Abaixo, cada uma dessas métricas é explicada em detalhes:

##### 4.6.2.1 GANHO DE INFORMAÇÃO (*INFOGAIN*)

O Ganho de Informação mede a redução na entropia ao dividir um conjunto de dados baseado em um atributo específico. A entropia é uma medida da incerteza ou desordem no conjunto de dados. Essa métrica ajuda a identificar atributos que mais reduzem a incerteza. O ganho de informação é calculado pela diferença entre a entropia do conjunto de dados original e a entropia após a divisão:

$$\text{InfoGain}(A) = \text{Entropia}(D) - \sum_{v \in \text{Valores}(A)} \frac{|D_v|}{|D|} \times \text{Entropia}(D_v)$$

#### Equação 7. Fórmula para calcular o InfoGain.

Onde:

- **Entropia** ( $D$ ) é a entropia do conjunto de dados  $D$ .
- $\text{Valores}(A)$  são os valores possíveis do atributo  $A$ .
- $D_v$  é o subconjunto de  $D$  onde  $A$  tem valor  $v$ .

#### 4.6.2.2 TAXA DE GANHO (*GainRatio*)

A Taxa de Ganho é uma variação do ganho de informação que normaliza o valor considerando o número e o tamanho dos ramos resultantes. Esta métrica ajusta o ganho de informação penalizando os atributos com muitos valores, o que pode evitar o sobreajuste.

$$\text{GainRatio}(A) = \frac{\text{InfoGain}(A)}{\text{SplitInfo}(A)}$$

#### Equação 8. Fórmula para calcular o GainRatio

Onde:

- $\text{SplitInfo}(A)$  é a informação gerada pela divisão do conjunto de dados, baseada no atributo  $A$ .

#### 4.6.2.3 ÍNDICE GINI

O Índice Gini proporciona uma medida direta da impureza, sendo útil para entender a homogeneidade das divisões na árvore. A impureza é calculada como a probabilidade de uma amostra ser classificada incorretamente se fosse rotulada aleatoriamente de acordo com a distribuição das classes no subconjunto. O Índice Gini é calculado da seguinte forma:

$$\text{Gini}(D) = 1 - \sum_{i=1}^n p_i^2$$

#### Equação 9. Fórmula para calcular o Índice Gini.

#### 4.6.2.4 QUI-QUADRADO ( $\chi^2$ )

A métrica Qui-Quadrado avalia a independência entre uma variável e a classe de destino. Essa métrica fornece uma visão estatística da independência, essencial para identificar relações significativas entre variáveis e a classe de destino. É calculada com base na diferença entre os valores observados e esperados em uma tabela de contingência. A fórmula é:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

**Equação 10. Fórmula para calcular a métrica Qui-Quadrado.**

Onde  $O_i$  são os valores observados e  $E_i$  são os valores esperados.

#### 4.7 INTEGRAÇÃO E AMOSTRAGEM DE DADOS

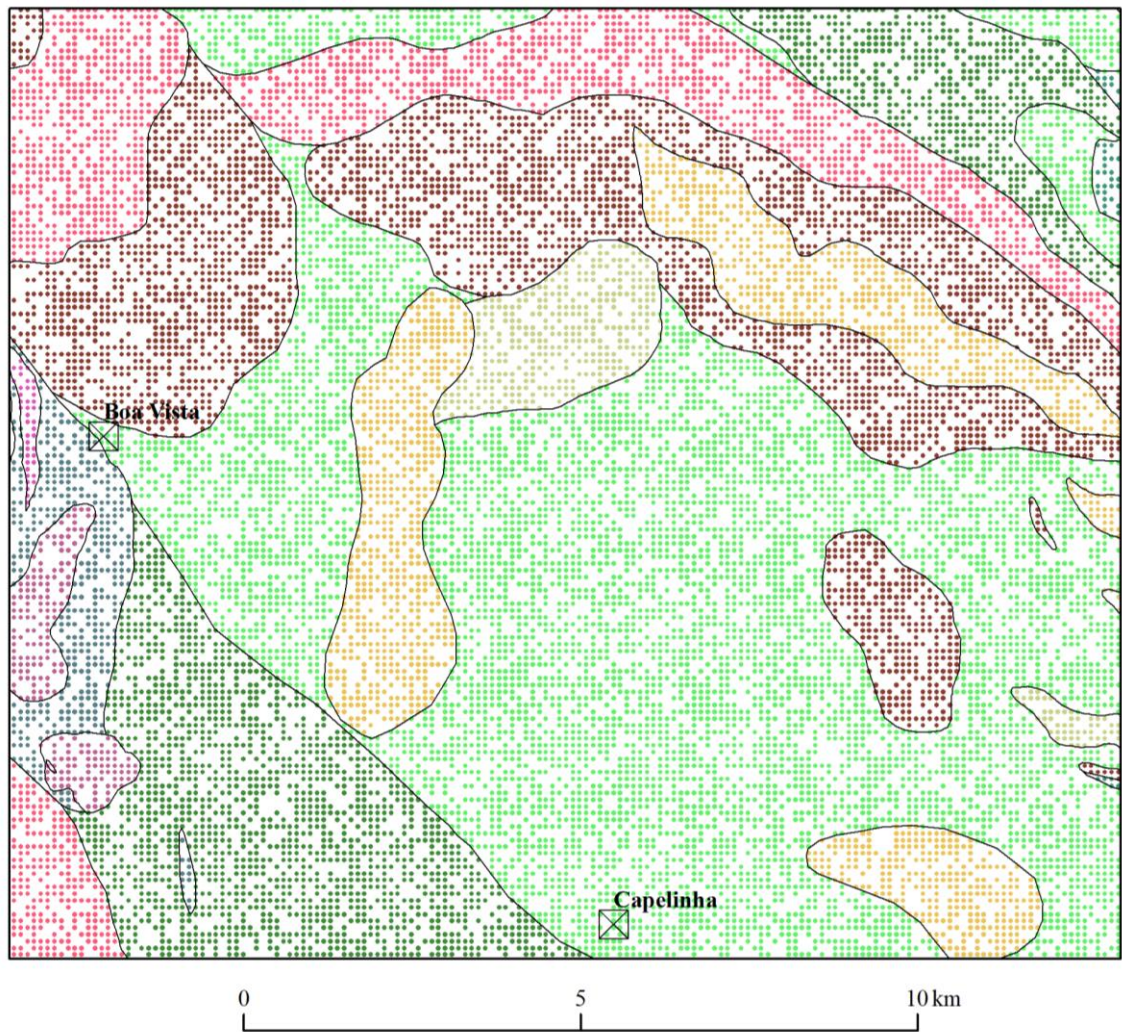
Para realizar o treinamento do modelo, os dados disponíveis foram integrados e amostrados utilizando-se ferramentas de diferentes *softwares*:

- **QGIS:** Utilizado para criar a malha de pontos com 100m de distância entre os pontos, utilizando a ferramenta Malha de Pontos.
- **Oasis Montaj:** Utilizado para integrar e filtrar os dados geofísicos, além de fornecer os valores dos *grids* para a malha de pontos gerada pelo QGIS.
- **ArcMap:** Utilizado para amostrar os dados através das ferramentas *Sample* e *Spatial Join*.

O arquivo *shapefile* (SHP) resultante foi exportado como um arquivo *Comma Separated Value* (CSV), que serve como a base para as análises subsequentes. No total foram utilizados 103.653 pontos, dos quais 66% (68.411) foram utilizados para treinar o modelo (Tabela 4). A densidade dos pontos ao longo da área pode ser observada na Figura 5.

**Tabela 4: Variáveis utilizadas para treinamento do modelo, com indicação de tipo (variáveis originais, variáveis derivadas de outras variáveis).**

<b>Sigla</b>	<b>Unidade/Grupo</b>	<b>Pontos Amostrados</b>	<b>Pontos Totais</b>	<b>Proporção Amostra/Total (%)</b>
N1dl	Coberturas detrito-lateríticas ferruginosas	212	7330	2,89%
NP1 $\gamma$ cl	Granito Córrego da Lavrinha	3	106	2,83%
NP1 $\gamma$ ms	Metagranito Sanclerlândia	339	11698	2,90%
NP1 $\gamma$ sj	São João	5	177	2,82%
NP1aibu	Anicums-Itaberaí, básico- ultrabásica	249	8600	2,90%
NP1aibutx	Anicums-Itaberaí, talco xisto	0	15	0,00%
NP1aipp	Anicums-Itaberaí, psamo-pelítica	166	5731	2,90%
NP1aippmx	Anicums-Itaberaí, mica-xisto	94	3247	2,89%
NP1aippqt	Anicums-Itaberaí, quartzito	2	82	2,44%
NP1airch	Anicums-Itaberaí, metachert	25	866	2,89%
NP1aircmm	Anicums-Itaberaí, mármore	37	1283	2,88%
NP1aircr	Anicums-Itaberaí, metarritmito	24	840	2,86%
NP1aito	Anápolis-Itaçu, ortogranulito	388	13397	2,90%
NP1aitoa	Anápolis-Itaçu, granulito ácido	297	10245	2,90%
NP1aitob	Anápolis-Itaçu, granulito básico	106	3644	2,91%
NP1aitp	Anápolis-Itaçu, paragranulito	222	7679	2,89%
NP2 $\gamma$ c	Granito Choupana	89	3089	2,88%
NP2cam	Campestre	52	1808	2,88%
NP2camg	Campestre, metagranito	234	8079	2,90%
Corpos M-UM	Córrego Seco, metagabro	233	8046	2,90%
NP23 $\gamma$ t	Metagranito Turvânia	52	1802	2,89%
NP $\gamma$ l	Granito Lage	87	2992	2,91%
NP $\gamma$ sr	Metadiorito Córrego Santa Rosa	48	1654	2,90%
PP4MP1mos	Mossâmedes, sedimentar	36	1243	2,90%
<b>Total</b>		<b>3000</b>	<b>103653</b>	<b>2,89%</b>



**Figura 5.** Densidade dos pontos usados para treinamento ao longo da área de pesquisa. As cores são simbólicas e variam de acordo com a litologia amostrada.

A amostragem foi realizada de forma estratificada, de modo que a quantidade de pontos varia de acordo com o tamanho da amostra, mas a proporção se mantém. Essa divisão equilibra a quantidade de dados para treinamento e teste, reduzindo o risco de *overfitting* enquanto ainda fornece uma base sólida para avaliar a generalização do modelo.

Gholamy, Kreinovich e Kosheleva (2018) explicam que a proporção 70/30 é frequentemente utilizada porque empiricamente resulta em uma boa performance do modelo. Isso proporciona um balanço entre a quantidade de dados para ajuste dos parâmetros do modelo e dados suficientes para validação e teste, assegurando que o modelo não apenas memorize os dados de treino, mas também se comporte bem com dados novos e não vistos.

Para o modelo deste trabalho foi utilizada a proporção 2/3 para amostragem e 1/3 para teste. Embora a proporção da amostragem siga o modelo geral de algoritmos de *ML*, a aplicação do método foi embasada em trabalhos envolvendo uso desses algoritmos para fins geológicos (Tabosa, 2023; Costa *et al.*, 2019; Kuhn *et al.*, 2019).

#### **4.8 TREINAMENTO E APLICAÇÃO DO MODELO *RANDOM FOREST***

A escolha do *Random Forest* como algoritmo para a classificação foi fundamentada em sua capacidade de lidar com dados complexos e heterogêneos, além de sua robustez contra o *overfitting*. O *workflow* foi conduzido no *software Orange Data Mining*, uma plataforma visual gratuita de análise de dados.

##### **Passos para a execução do *Random Forest*:**

###### **1. Preparação dos Dados:**

- Dados geológicos, geofísicos e de sensoriamento remoto foram integrados em um único arquivo *CSV*.
- A normalização *Z-score* foi aplicada para ajustar os valores das variáveis.
- A imputação de valores ausentes foi realizada utilizando a mediana dos valores.

###### **2. Engenharia de Recursos:**

- As variáveis derivadas foram criadas a partir dos dados geofísicos originais.

###### **3. Configuração do Modelo:**

- A validação cruzada *k-fold* ( $k = 3$ ) foi utilizada para avaliar a performance do modelo, minimizando o risco de *overfitting*.
- Hiperparâmetros foram otimizados utilizando o *Randomized Search*, ajustando o número de árvores, a profundidade máxima das árvores e o número mínimo de amostras por folha.

###### **4. Treinamento e Avaliação:**

- O modelo foi treinado com uma amostra estratificada de 3.000 pontos, representando 2,89% do total de 103.653 pontos.
- As métricas de desempenho (acurácia, precisão, *recall*, *F1-score*, matriz de confusão) foram calculadas para avaliar a eficácia do modelo.

#### **4.9 CLASSIFICAÇÃO DAS AMOSTRAS DE CAMPO**

Durante e após a etapa de campo, as amostras coletadas foram descritas de forma macroscópica, sendo classificadas de acordo o sistema proposto por Streckeisen (1976). Esse trabalho foi realizado com auxílio de lupa de mão, lupa de bancada e instrumentos para identificar as propriedades físicas dos minerais (lupa, ímã, canivete, ácido clorídrico, minerais da escala de dureza, dentre outros).

## 5. RESULTADOS E DISCUSSÕES

Neste capítulo, serão apresentados os resultados obtidos a partir da aplicação do modelo treinado do algoritmo *Random Forest* para a identificação de corpos máficos-ultramáficos assim como as discussões sobre os mesmos. A estrutura do capítulo inclui a avaliação do modelo, previsões geradas para corpos máficos-ultramáficos e validação dos resultados com dados de campo. Ao longo dos tópicos são apresentados tabelas, mapas e visualizações que ilustram os resultados.

### 5.1 AVALIAÇÃO DO MODELO

Nesta seção, é apresentada uma análise do desempenho do modelo *Random Forest* utilizado para a identificação e caracterização de corpos máficos-ultramáficos na área de estudo. A avaliação é feita com base em métricas de desempenho, importância das variáveis e a matriz de confusão, proporcionando uma visão abrangente sobre a eficácia do modelo aplicado.

#### 5.1.1 MÉTRICAS DE DESEMPENHO

As principais métricas de desempenho do modelo para todas as litologias e especificamente para os corpos máficos-ultramáficos são apresentadas na Tabela 5.

**Tabela 5: Métricas de desempenho do algoritmo *Random Forest*. Acurácia; F1-Score; Precisão; Recall; Matthew Correlation Coefficient (MCC).**

<b>Métrica</b>	<b>Predições Gerais</b>	<b>Corpos M-UM</b>
<b>CA</b>	0.827	0.981
<b>F1-Score</b>	0.827	0.874
<b>Precisão</b>	0.829	0.905
<b>Recall</b>	0.827	0.845
<b>MCC</b>	0.812	0.864

A análise dessas métricas mostra que o modelo demonstrou um desempenho de bom a muito bom em suas predições, com uma acurácia geral de 82,7%, um *FI-Score* de 82,7%, uma precisão de 82,9%, um *Recall* de 82,7% e um Coeficiente de Correlação de Matthews (MCC) de 81,2%. Para a classe específica dos corpos máficos-ultramáficos (M-UM), o modelo também apresentou resultados robustos, com uma acurácia de 98,1%, um *FI-Score* de 87,4%, uma precisão de 90,5%, um *Recall* de 84,5% e um MCC de 86,4%.

Tendo em vista que 50% proporcionam acertos ao acaso e 100% representam um modelo perfeito, esses resultados indicam que o modelo adotado neste trabalho é eficaz e confiável tanto de forma geral quanto na classificação específica dos corpos M-UM.

### 5.1.2 IMPORTÂNCIA DAS VARIÁVEIS

A tabela 6 fornece uma análise detalhada da importância das variáveis utilizadas no modelo de Random Forest. Para cada variável, são exibidas quatro métricas principais: ganho de informação, taxa de ganho, índice Gini e  $\chi^2$  (Qui-quadrado). As bandas espectrais (variáveis B1 a B7) não desempenharam um papel significativo neste modelo, obtendo métricas inferiores ao IDD.

**Tabela 6: Ranking de importância das variáveis utilizadas com base no Ganho de Informação, Taxa de Ganho, Índice Gini e Qui-quadrado ( $\chi^2$ ).**

Variável	Ganho de Informação	Taxa de Ganho	Gini	$\chi^2$
eTh	0,395	0,197	0,035	32413,105
eU	0,303	0,151	0,029	28842,222
eTh/K	0,290	0,145	0,023	25564,852
K	0,286	0,143	0,022	25904,526
eU/K	0,263	0,132	0,022	20336,043
ASA/eTh	0,257	0,128	0,024	21515,740
eU/eTh	0,243	0,121	0,027	19854,591
SRTM	0,240	0,120	0,017	17850,447
ASA/K	0,226	0,113	0,019	19962,293
ASA	0,163	0,082	0,015	12960,022
ASA/eU	0,145	0,072	0,013	6299,917
IDD	0,058	0,029	0,004	4968,813
B5	0,026	0,013	0,002	2410,357
B6	0,017	0,008	0,001	1006,588
B7	0,016	0,008	0,001	1042,216

A contagem de eTh destaca-se como a variável mais informativa e relevante, liderando em todas as métricas com um Ganho de Informação de 0,395, uma Taxa de Ganho de 0,197, um índice de Gini de 0,035 e um valor  $\chi^2$  de 32413,105. Estes valores indicam que a contagem de eTh é crucial para o desempenho do modelo, fornecendo a maior quantidade de informação útil para a classificação das litologias.

Em seguida, a contagem de eU também se mostra altamente relevante, com um Ganho de Informação de 0,303, uma Taxa de Ganho de 0,151, um índice de Gini de 0,029 e um valor  $\chi^2$  de 28842,222. Isso reflete sua importância significativa no modelo.

A razão eTh/K% e a porcentagem de K% são igualmente notáveis, apresentando valores elevados em todas as métricas. A razão eTh/K% possui um Ganho de Informação de 0,29, uma Taxa de Ganho de 0,145, um índice de Gini de 0,023 e um valor  $\chi^2$  de 25564,852. Enquanto a porcentagem de K% tem um Ganho de Informação de 0,286, uma Taxa de Ganho de 0,143, um índice de Gini de 0,022 e um valor  $\chi^2$  de 25904,526. Ambas são variáveis importantes, contribuindo de maneira significativa para a precisão do modelo.

Variáveis como as razões eU/K%, ASA/eTh e eU/eTh, também demonstram importância considerável, com Ganhos de Informação de 0,263, 0,257 e 0,243 respectivamente, além de valores consistentes nas outras métricas. Estas variáveis contribuem para a discriminação eficaz entre as classes.

Por outro lado, variáveis como "ASA", "ASA/eU" e "IDD" apresentam menor relevância comparativa, com valores de Ganho de Informação de 0,163, 0,145 e 0,058 respectivamente. Seus menores valores de Taxa de Ganho, Gini e  $\chi^2$  indicam que têm um impacto menos significativo nas previsões do modelo.

Em resumo, as variáveis "eTh", "eU", "eTh/K%" e "K%" são identificadas como as mais importantes para o desempenho do modelo gerado, fornecendo maior quantidade de informação útil e contribuindo significativamente para a precisão e confiabilidade das previsões. Outras variáveis também desempenham papéis importantes, mas em menor grau.



é um classificador eficaz, embora ainda existam algumas confusões que podem ser reduzidas com ajustes no modelo, ou com rótulos advindos de um mapeamento de maior detalhe.

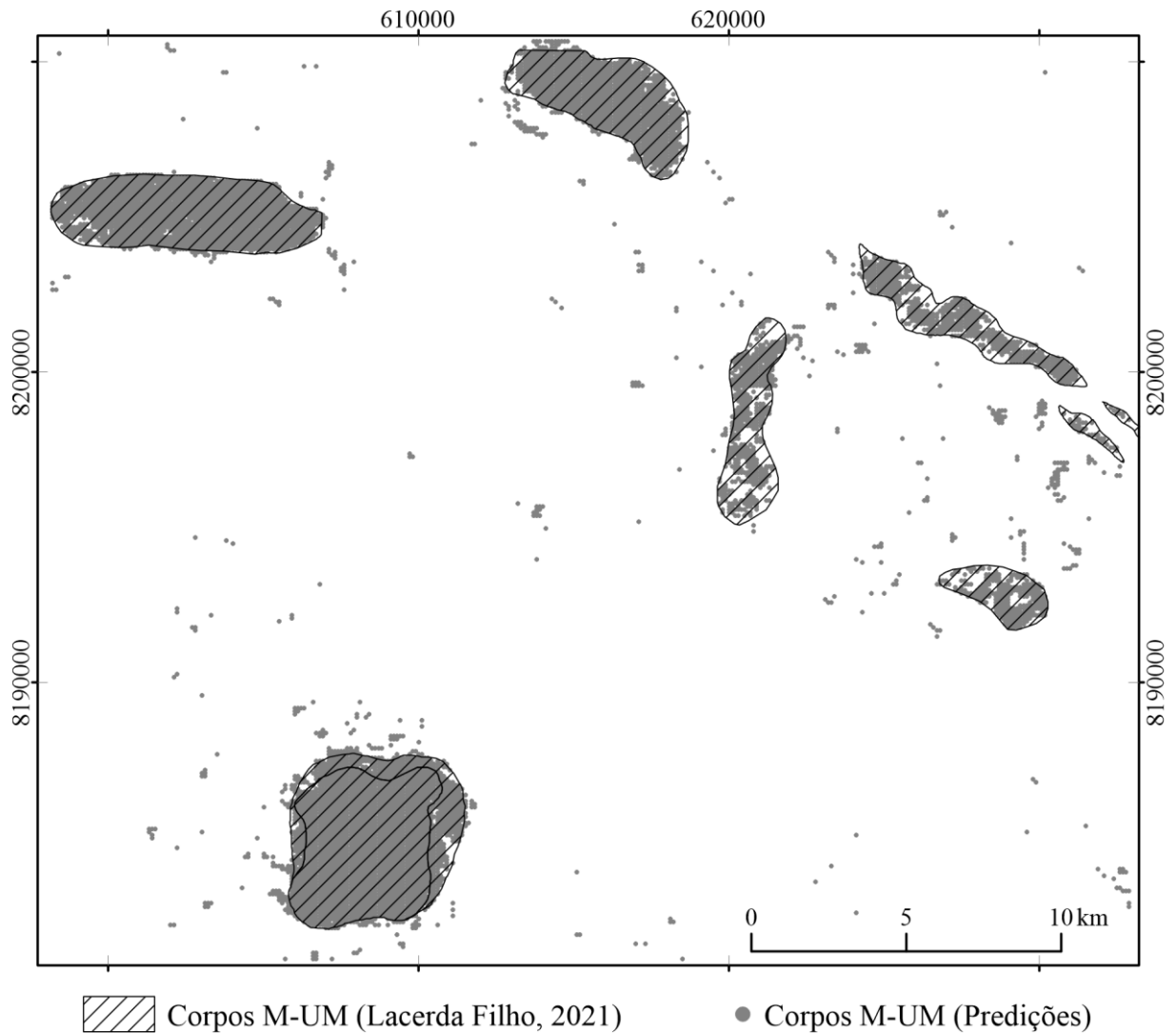
Vale mencionar que as classes NP1 $\gamma$ cl e NP1aito também promoveram falsos positivos, 5% e 4% respectivamente, associando seus parâmetros a pontos onde estão mapeadas litologias da classe dos corpos M-UM. Isso realça novamente a interação da unidade NP1aito com os corpos M-UM e traz à tona a região de proximidade entre o Granito Córrego da Lavrinha (NP1 $\gamma$ cl) e o Corpo Córrego Seco (NP3 $\delta$ asc).

## **5.2 PREDIÇÕES DO MODELO**

Após a reclassificação das unidades alvo para formar a classe Corpos M-U, foi obtido um conjunto de litotipos composto por: gabros, metagabros, piroxenitos, peridotitos, metaperidotitos, dioritos, metadioritos, noritos, metanoritos, gabronoritos, hornblenda gabros, hornblenda gabronoritos, xitos máficos e anfibolitos. Dito isto, o modelo preditivo foi preparado para identificar e mapear esses litotipos, ou litotipos com características similares. Nesse tópico serão discutidos os resultados obtidos.

### **5.2.1 MAPA PREDITIVO PARA CORPOS MÁFICOS-ULTRAMÁFICOS**

O modelo treinado de *Random Forest* foi aplicado com êxito para identificar corpos máficos-ultramáficos na área de estudo. Os resultados das classificações foram filtrados para a classe de Corpos M-UM, resultando em um mapa preditivo (Figura 6). As predições geradas pelo modelo citado foram comparadas com dados geológicos conhecidos e amostras de campo para validar a eficácia e veracidade das identificações.

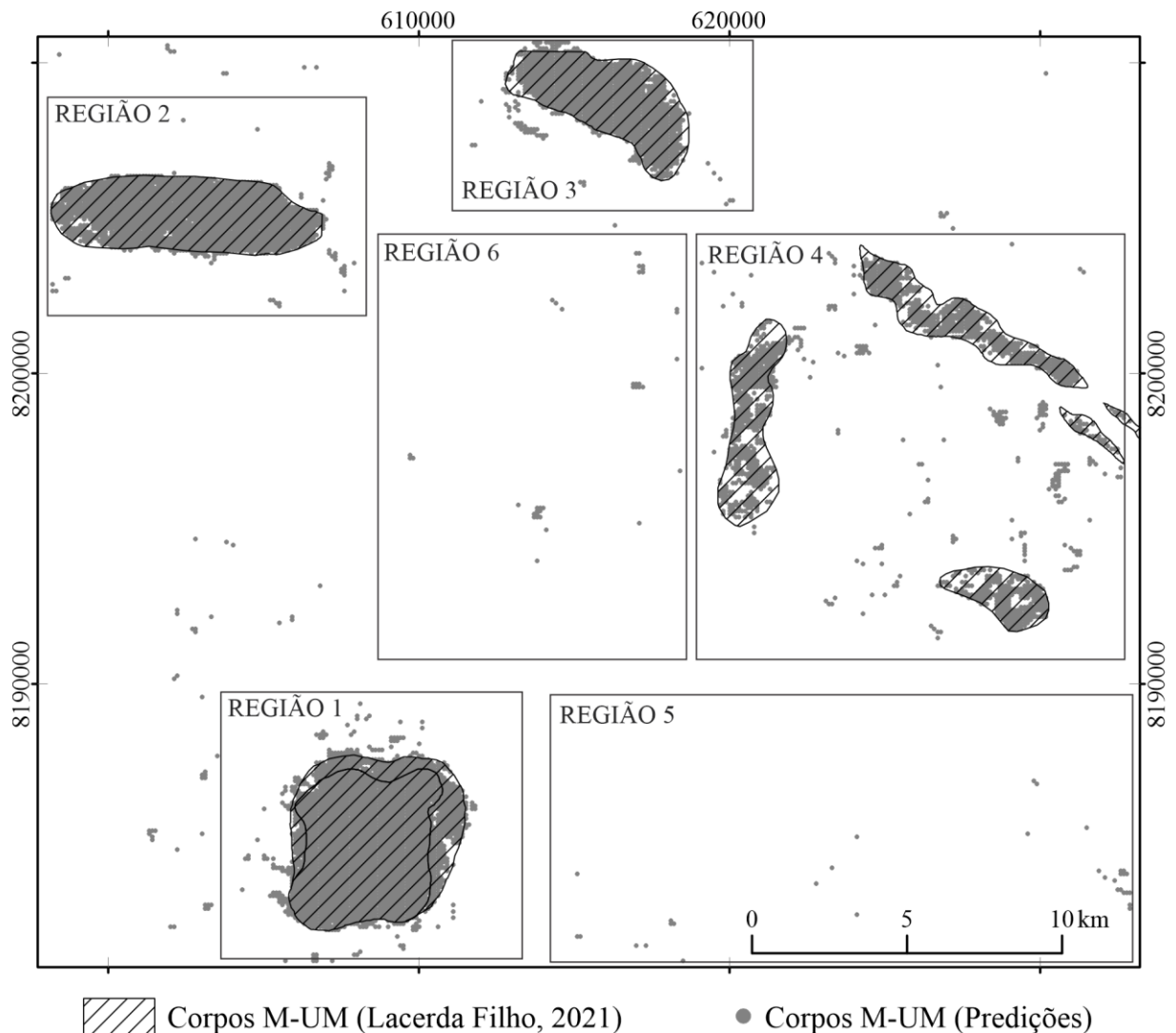


**Figura 6. Mapa preditivo gerado pelo modelo Random Forest indicando a classe de Corpos Máficos-Ultramáficos na área de estudo.**

De um modo geral, o resultado das predições foi fidedigno às delimitações fornecidas para os corpos M-UM.

### 5.3 VALIDAÇÃO DOS DADOS EM CAMPO

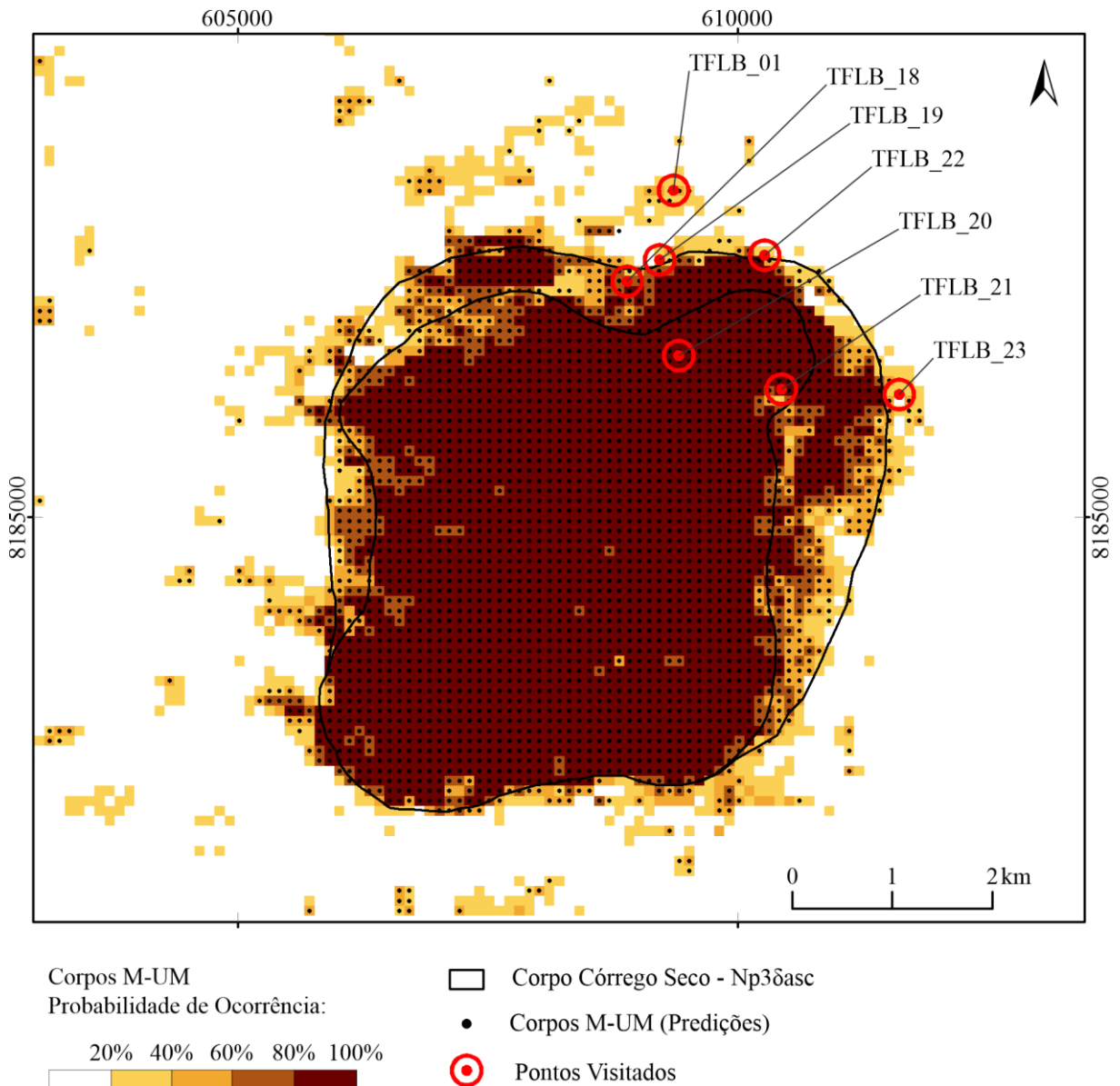
Para discorrer sobre os resultados com maior fluidez o mapa preditivo foi subdividido em 6 regiões, englobando os corpos M-UM e os pontos de visitação (Figura 7). Devido a escala do trabalho, os pontos visitados foram definidos de forma estratégica, levando em consideração cumulados de pontos de predições e a facilidade de acesso.



**Figura 7.** Mapa preditivo gerado pelo modelo Random Forest, indicando a distribuição predita de corpos máficos-ultramáficos na área de estudo. As subdivisões regionais para a apresentação dos resultados estão destacadas em polígonos numerados. Região 1: Corpo Córrego Seco; Região 2: Americano do Brasil; Região 3: Gongomé-Retiro; Região 4: Águas Claras de Goiás, Capelinha dos Gomes e Fazenda Paraíso; Regiões 5 e 6: Predições isoladas.

### 5.3.1 REGIÃO 1: SW DA ÁREA DE ESTUDO - CORPO CÓRREGO SECO

As predições geradas para a região 1 (Figura 8), que compreende o entorno do corpo intrusivo Córrego Seco (NP3δasc), localizado na porção sudoeste da área de estudo, foram satisfatórias. Os pontos de predição delineiam bem a sua delimitação original, o que pode indicar que os contatos traçados advêm de um contraste bem definido nos produtos geofísicos.



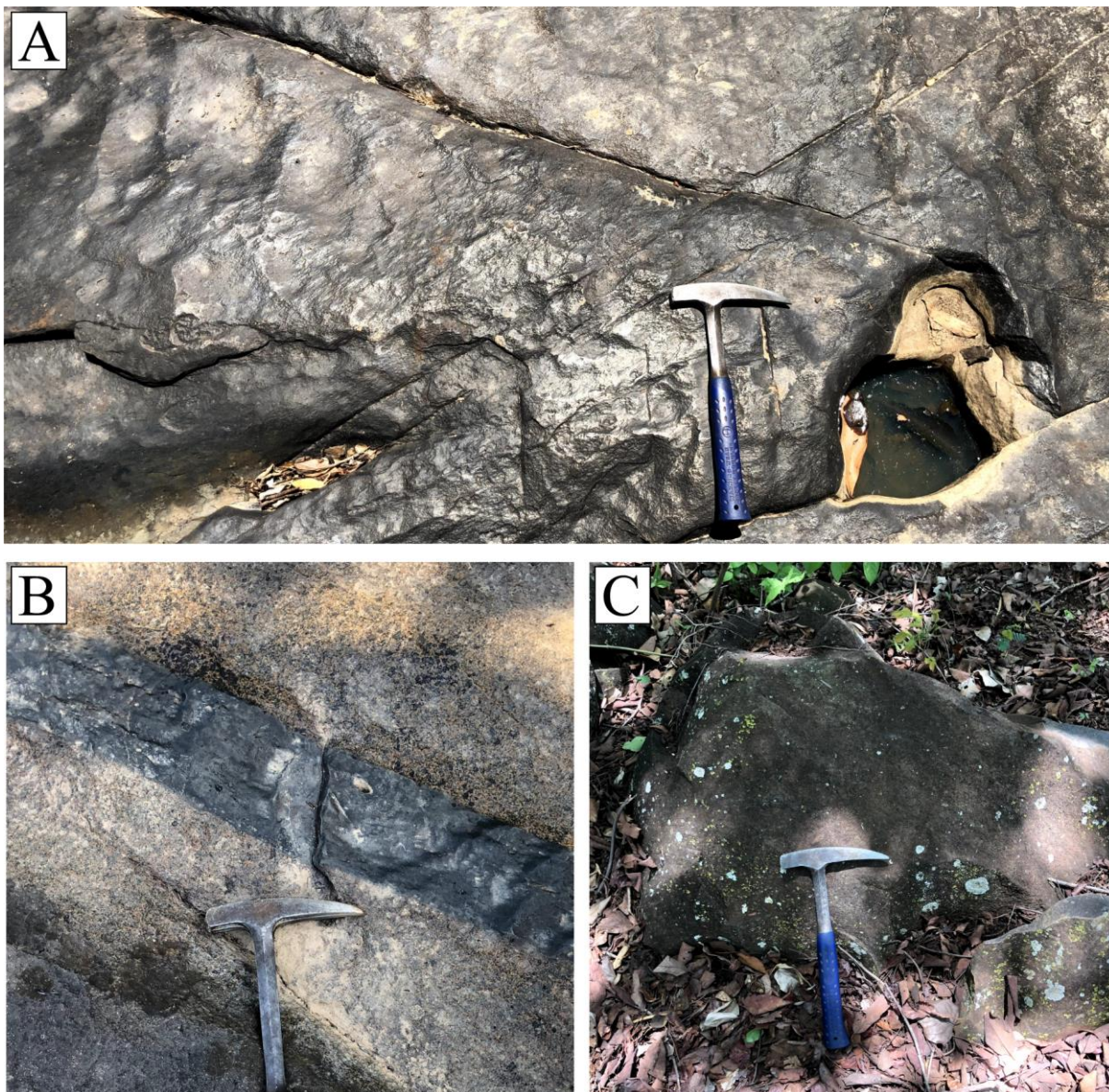
**Figura 8. Mapa preditivo com enfoque na região do Corpo Córrego Seco (NP3δasc). O mapa foi gerado pelo modelo Random Forest e indica a probabilidade de ocorrência e os pontos de predição da classe de Corpos Máficos-Ultramáficos na área de estudo.**

No entanto, ainda é possível observar um pequeno estreitamento das bordas, o que pode indicar respostas advindas das rochas encaixantes, refinando essa interação. Importante realçar que as predições também formaram pequenos cumulados nas extremidades do corpo, em especial nas porções noroeste, oeste-sudoeste e sul-sudeste do corpo.

Nessa região, foram realizados 7 pontos de visita abrangendo a porção NE do corpo. Ao longo dos pontos visitados, os afloramentos encontrados em sua maioria são em forma de lajedo ou blocos/matacões desarticulados.

O ponto TFLB\_01 está localizado fora da delimitação do Corpo Córrego Seco, em um cumulado de pontos de predição e em zonas de baixa a razoável (40%) de probabilidade de ocorrência da classe Corpos M-UM. Nesta porção não foram encontrados afloramentos, é uma área utilizada para plantio. O solo possui coloração avermelhada e granulação silto-argilosa, descrição típica de solos formados a partir de rochas ferromagnesianas. Pensa-se que a resposta obtida para esse ponto, pode ser advinda de subsuperfície, captada pelo produto magnetométrico.

Os pontos TFLB\_20 (Figuras 9A e 9B), TFLB\_21 e TFLB\_22 (Figura 9C) estão localizados dentro da delimitação do Corpo Córrego Seco, em pontos de predição e em zonas de altíssima probabilidade (80-100%) de ocorrência da classe Corpos M-UM. Nestes pontos foram observadas rochas gabróicas de granulação média a grossa, magnetismo moderado e compostas principalmente por piroxênio e plagioclásio, alternando entre uma predominância do plagioclásio e maior teor de minerais máficos em diferentes pontos dos afloramentos. Em alguns pontos os gabros também são cortados por diques de diabásio com direção NW-SE (Figura 9B).



**Figura 9. (A) Afloramento de rochas gabróicas no leito do Rio dos Bois, com padrão de fraturamento em “X” nas direções NNE-SSW e ENE-WNW (Ponto: TFLB\_20). (B) Gabro cortado por dique de diabásio com direção NW-SE. (Ponto TFLB\_20). (C) Blocos de gabro de granulação média (TFLB\_22).**

Os pontos TFLB\_18, TFLB\_19 e TFLB\_23 estão localizados nos limites da delimitação do Corpo Córrego Seco - Fácies Dioríticas (NP3δascdi), em pontos de predição e em zonas de média à alta probabilidade (60-80%) de ocorrência da classe Corpos M-UM. Os afloramentos se caracterizam como grandes lajedos (Figura 10A). Nestes pontos foram observadas rochas dioríticas e granodioríticas (Figura 10), mesocráticas, faneríticas a subfaneríticas, fracamente magnéticas, de granulação média, e compostas principalmente por piroxênio (dioritos), plagioclásio, biotita e quartzo (nos granodioritos ~20%).

As principais feições ígneas observadas nos afloramentos destas rochas remetem a misturas física (*mingling*) entre magma máfico e félsico, dada pela assimilação de fenocristais de feldspato da porção félsica presentes na porção máfica (Figura 10 B).

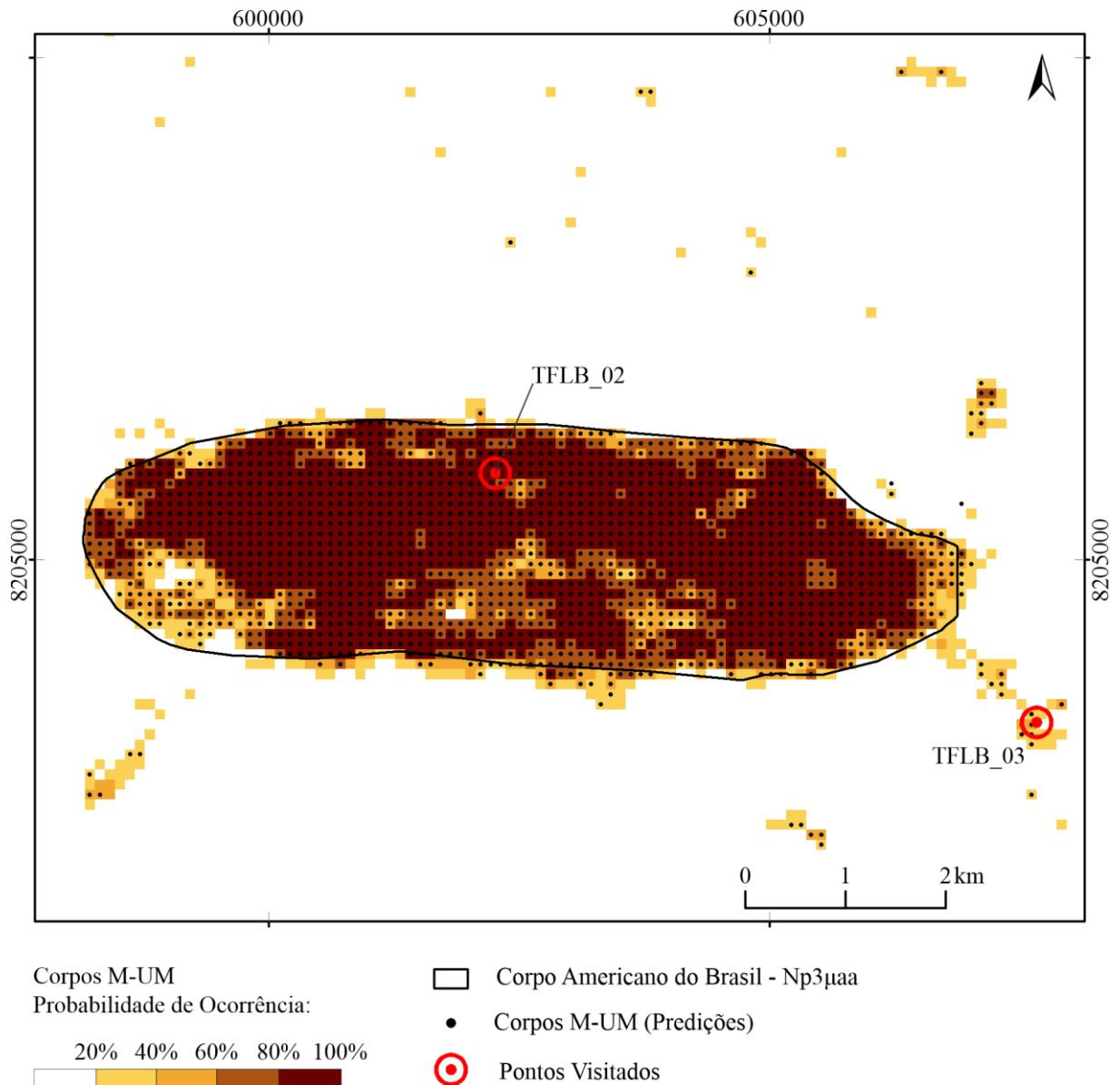


**Figura 10. (A) Fotografia aérea tirada através de drone. Rocha de composição granodiorítica disposta em lajedo. (Ponto: TFLB\_19). (B) Mistura física entre fenocristais de feldspato de rocha félsica em rocha máfica. (Ponto TFLB\_19).**

Tendo em vista a localização dos pontos de predição, os litotipos encontrados e as probabilidades de ocorrência, pode-se concluir que o modelo obteve um resultado satisfatório nesta região. É perceptível a variação dos litotipos de acordo com o grau de confiança das predições, isto é, nas zonas de maior probabilidade foram encontradas rochas máficas melanocráticas, enquanto nas zonas de probabilidade razoável, rochas mesocráticas com variações faciológicas.

### **5.3.2 REGIÃO 2: NW DA ÁREA DE ESTUDO - CORPO AMERICANO DO BRASIL**

Na região 2 (Figura 11), localizada na porção noroeste da área, que inclui o complexo Americano do Brasil (NP3 $\mu$ aa), as predições também foram satisfatórias.



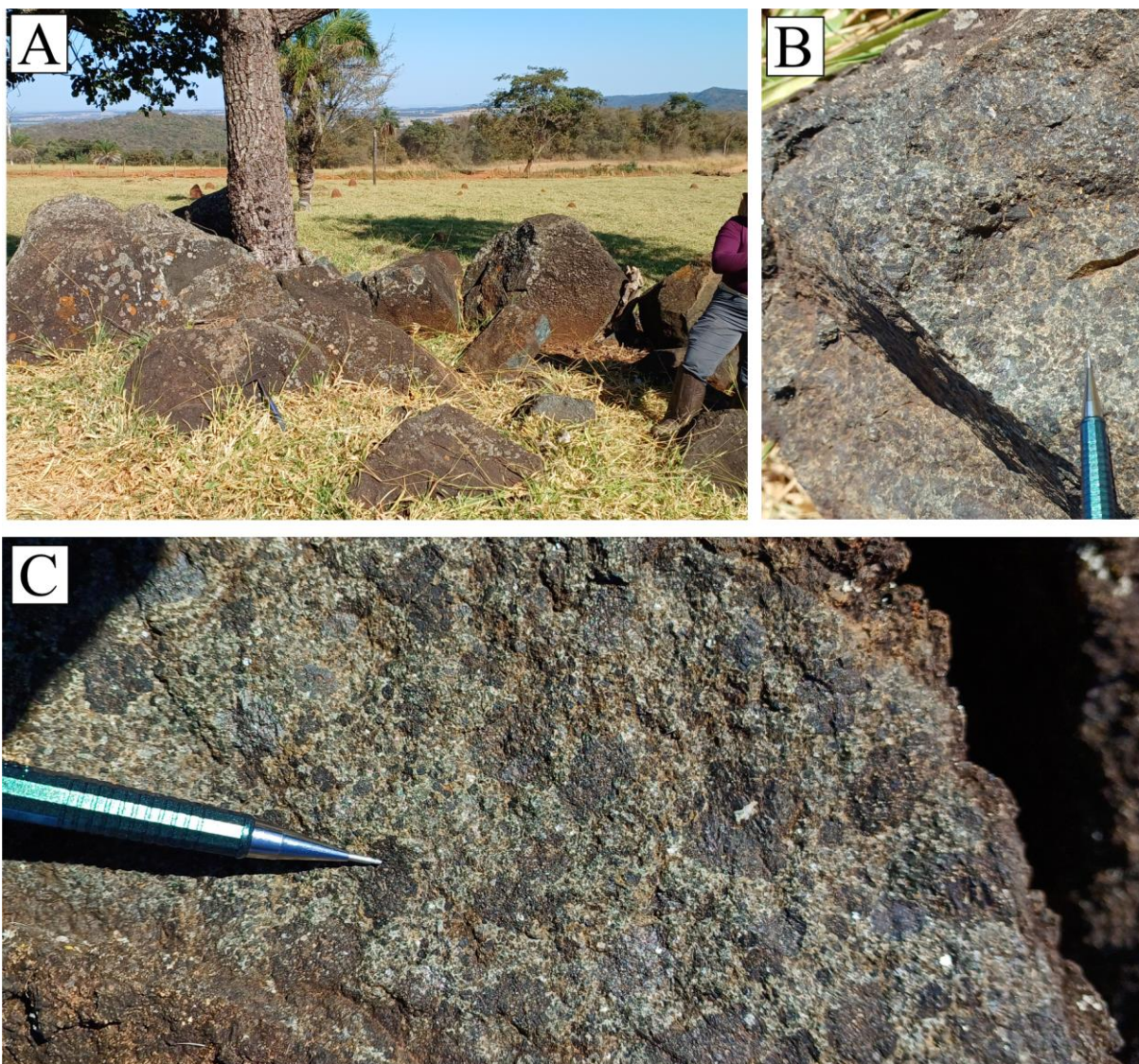
**Figura 11. Mapa preditivo com enfoque na região do Complexo Americano do Brasil (NP3μaa). O mapa foi gerado pelo modelo Random Forest e indica a probabilidade de ocorrência e os pontos de predição da classe de Corpos Máficos-Ultramáficos na área de estudo.**

De modo geral, os pontos de predição indicaram corretamente a sua delimitação original. Isso pode indicar que, como no tópico anterior, os contatos traçados podem ter sido embasados em um contraste geofísico bem definido. Dito isto, vale observar a formação de “vazios” e zonas de baixa probabilidade na porção sudoeste do corpo.

Nessa região, foram realizados 2 pontos de visitação abrangendo a porção central do corpo e um dos cumulados de pontos preditivos.

O ponto TFLB\_03 está localizado fora da delimitação do Corpo Americano do Brasil, em um acúmulo de pontos preditivos e em zonas de baixa probabilidade (40%) de ocorrência da classe Corpos M-UM. Nesta porção, assim como no ponto TFLB\_01, não foram encontrados afloramentos, também sendo uma área utilizada para plantio. O solo possui coloração marrom claro e granulação argilo-siltosa. Pensa-se que a resposta obtida para essa porção, pode ter sido dada pela baixa probabilidade de ser qualquer outra litologia. Uma vez que o rótulo preditivo é uma votação dentre as probabilidades de cada um dos litotipos do modelo.

O ponto TFLB\_02 (Figura 12A) está localizado dentro da delimitação do Corpo Americano do Brasil, em pontos de predição e em zonas de altíssima probabilidade (80-100%) de ocorrência da classe Corpos M-UM. Neste ponto foram observadas rochas ultramáficas, maciças, inequigranular porfiríticas, de granulação média a grossa (Figura 12B), magnetismo moderado e compostas principalmente por piroxênio, plagioclásio, clorita e sulfetos em menor quantidade, com os piroxênios em processo de uralitização (Figura 12C).



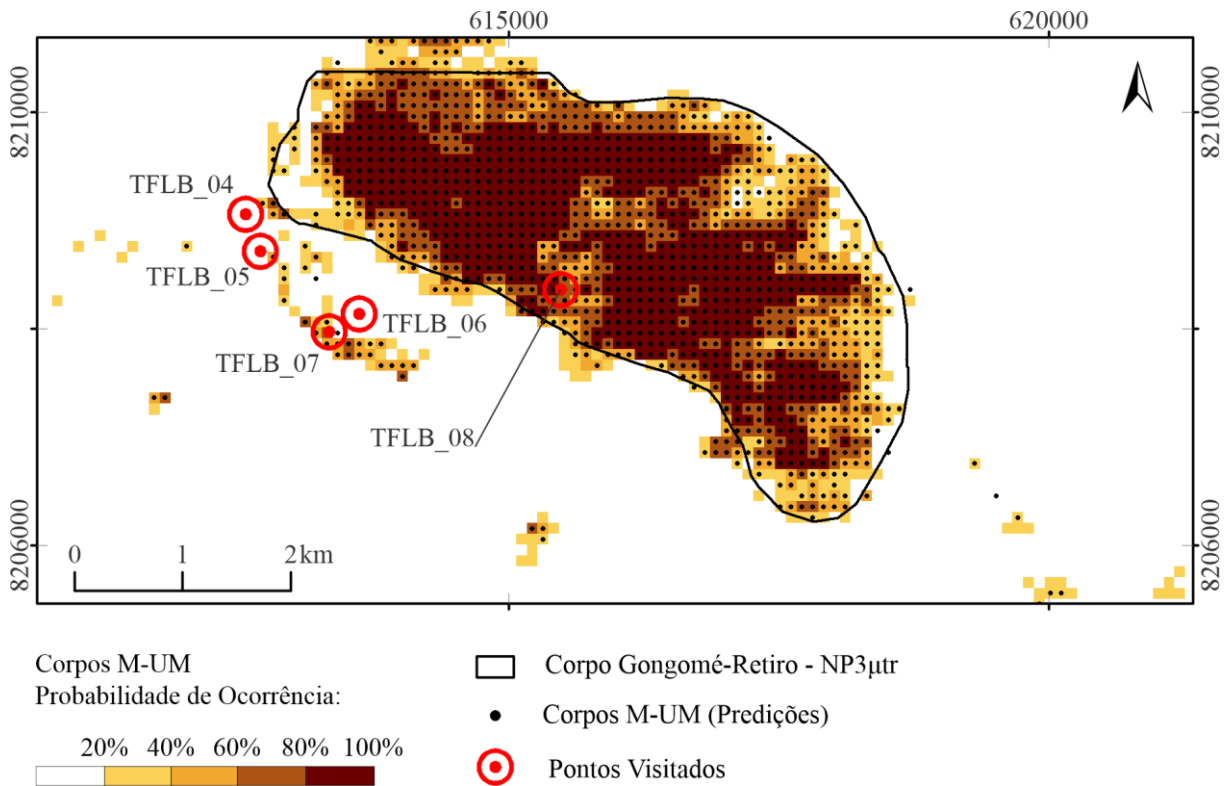
**Figura 12. (A) Afloramento de blocos e matacões de composição ultramáficas (Ponto: TFLB\_02). (B) Bloco de piroxenito exibindo textura inequigranular porfirítica e grau moderado de intemperismo (Ponto: TFLB\_02). (C) Amostra de piroxenito (Ponto: TFLB\_02).**

Tendo em vista a localização dos pontos de predição, os litotipos encontrados e as probabilidades de ocorrência, pode-se concluir que o modelo também obteve um resultado satisfatório na região 2. As predições seguiram de forma razoável a geometria original do corpo e obtiveram na sua zona de maior probabilidade rochas ultramáficas, enquanto nas zonas de probabilidade baixa houve ausência de afloramentos.

### **5.3.3 REGIÃO 3: N DA ÁREA DE ESTUDO - CORPO GONGOMÉ-RETIRO**

A região 3 (Figura 13), localizada na porção norte da área de estudo e que engloba o corpo Gongomé-Retiro (NP3 $\mu$ tr), apresentou predições alinhadas com as observações de campo. A

delimitação dos corpos intrusivos foi confirmada pelas amostras coletadas, que revelaram rochas ferromagnesianas nas áreas previstas pelo modelo.



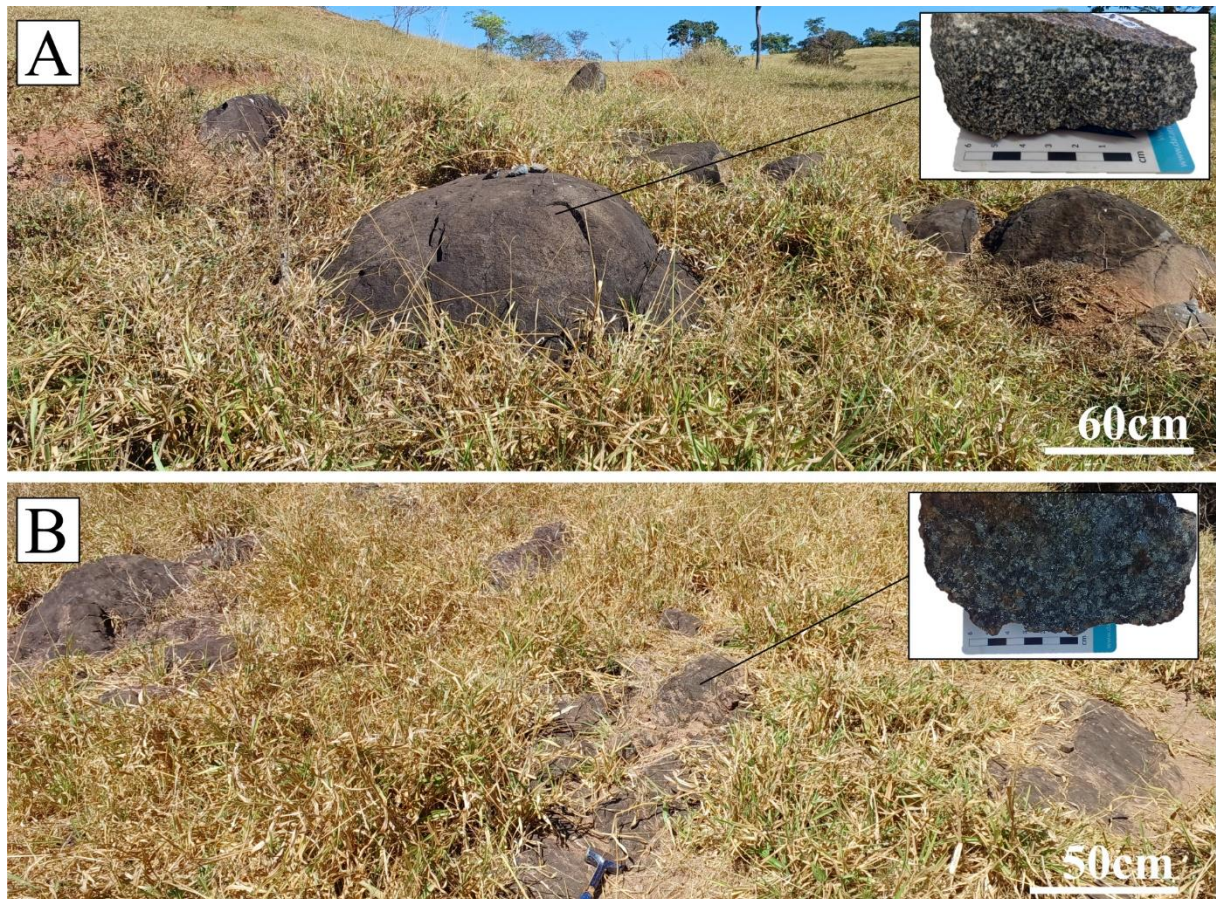
**Figura 13. Mapa preditivo com enfoque na região do Corpo Gongomé-Retiro (NP3μtr). O mapa foi gerado pelo modelo Random Forest e indica a probabilidade de ocorrência e os pontos de predição da classe de Corpos Máficos-Ultramáficos na área de estudo.**

Nessa região, foram realizados 5 pontos de visitaç o abrangendo a porç o centro-sul do corpo Gongom -Retiro e alguns acumulados de pontos fora da delimita o   sudoeste. Ao longo dos pontos visitados, os afloramentos encontrados em sua maioria est o dispostos em forma de blocos e matac es aut ctones, desarticulados, moderadamente intemperizados e pouco alterados.

Os pontos TFLB\_04, TFLB\_05, TFLB\_06 e TFLB\_07 est o localizados fora da delimita o do Corpo Gongom -Retiro, n o necessariamente acima de pontos de predi o e dentro de zonas de m dia a alta probabilidade (40-80%) de ocorr ncia da classe Corpos M-UM. Os pontos TFLB\_04, TFLB\_05 e TFLB\_06 foram escolhidos por serem locais muito pr ximos a pontos de predi o e com rochas aflorantes.

Os pontos TFLB\_04 (Figura 14A) e TFLB\_06 exibem gabros e quartzo-dioritos maciços, faner ticos de granula o m dia, inequigranulares, variando de melanocr ticos a

mesocráticos, compostos principalmente por piroxênio, plagioclásio, biotita, anfibólios, quartzo (~15-20% nos dioritos) e sulfetos disseminados (pirita, calcopirita, bornita). O ponto TFLB\_05 (Figura 14B) consiste em blocos desarticulados de piroxenito, rochas holomelanocráticas, de granulação média a grossa, textura inequigranular porfirítica, composta essencialmente por piroxênio, plagioclásio e minerais acessórios e de alteração em menor quantidade.



**Figura 14. (A) Afloramento de blocos e matacões de composição gabroica/diorítica (Ponto: TFLB\_04). (B) Afloramento de rochas ultramáficas na vizinhança sudoeste do Corpo Gongomé-Retiro (Ponto: TFLB\_05).**

Consequente, o ponto TFLB\_07 possui um contraste às rochas plutônicas e intermediárias, compreendendo um afloramento métrico de chl-tlc-anf-ms-xisto, mesocrático, bem intemperizado, com magnetismo fraco e relativamente denso.

O ponto TFLB\_08 (Figura 15A) está localizado dentro da delimitação do Corpo Gongomé-Retiro, em pontos de predição e em zonas de alta a altíssima probabilidade (60-100%) de ocorrência da classe Corpos M-UM. Neste ponto foram observados blocos e

matações gabróicos, maciços, inequigranular porfiríticos, de granulação média a grossa, magnetismo moderado e compostos principalmente por piroxênio, plagioclásio, granada e sulfetos em menor quantidade. Em superfícies intemperizadas é possível visualizar os cristais milimétricos de granada (Figura 15B).



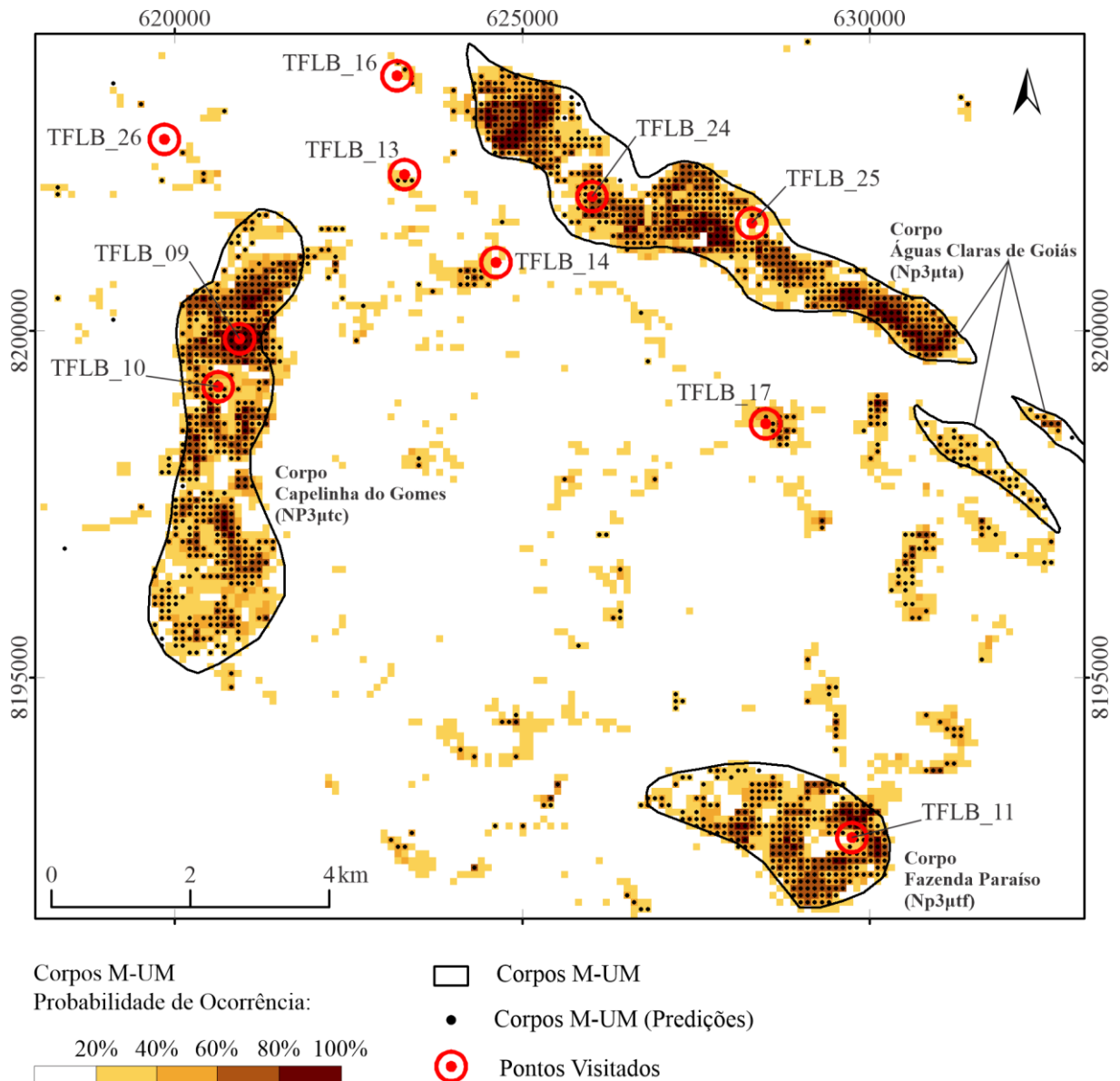
**Figura 15. (A) Blocos e matações de composição gabróica nas intermediações do Corpo Gongomé-Retiro (Ponto: TFLB\_08). (B) Detalhe da imagem “(A)” exibindo os cristais de granada em uma das superfícies intemperizadas, formando um aspecto de “catapora”.**

Analisando a localização dos pontos de predição, os litotipos encontrados e as probabilidades de ocorrência, pode-se concluir que o modelo obteve um bom resultado na região 3. Embora as predições não tenham se adequado a geometria original do corpo Gongomé-Retiro, os pontos de visitação obtiveram na sua zona de maior probabilidade rochas máficas a ultramáficas. E nas

zonas de probabilidade baixa, rochas máficas a intermediárias, incluindo os cumulos formados na vizinhança sudoeste do corpo.

#### **5.3.4 REGIÃO 4: E DA ÁREA DE ESTUDO - CORPOS ÁGUAS CLARAS DE GOIÁS, CAPELINHA DOS GOMES E FAZENDA PARAÍSO**

Na região 4, localizada na porção leste da área de estudo e que engloba os corpos: Águas Claras de Goiás (NP3 $\mu$ ta), Capelinha dos Gomes (NP3 $\mu$ tc) e Fazenda Paraíso (NP3 $\mu$ tf); o modelo obteve um bom desempenho, mas foi a região de maior divergência para com as delimitações originais (Figura 16). Os pontos preditivos ocuparam um espaço reduzido dentro das delimitações originais, reduzindo os limites ocupados por litotipos que tenham características comuns na classe de Corpos M-UM. Outro ponto que vale realçar foi a geração de uma quantidade razoável de acúmulos de predições fora desses limites, o que formou uma geometria semi-circular nesta região. As amostras de campo confirmaram a presença de litologias máficas a ultramáficas nas áreas preditas.



**Figura 16.** Mapa preditivo com enfoque na região dos corpos Águas Claras de Goiás, Fazenda Paraíso e Capelinha dos Gomes. O mapa foi gerado pelo modelo Random Forest e indica a probabilidade de ocorrência e os pontos de predição da classe de Corpos Máficos-Ultramáficos na área de estudo.

Nessa região, foram realizados 10 pontos de visitação abrangendo cada um dos corpos intrusivos principais e os acúmulos de predições mais acessíveis. Ao longo dos pontos visitados, os afloramentos encontrados em sua maioria são em forma de lajedo ou blocos/matacões desarticulados.

Os pontos TFLB\_24 e TFLB\_25 (Figura 17) estão localizados dentro da delimitação do Corpo Águas Claras de Goiás, em pontos de predição e em zonas de moderada a alta probabilidade (80-100%) de ocorrência da classe Corpos M-UM. Nestes pontos foram observadas rochas

metabásicas holomelanocráticas afaníticas e metagabros de granulação fina a média, compostos principalmente por piroxênio e plagioclásio.

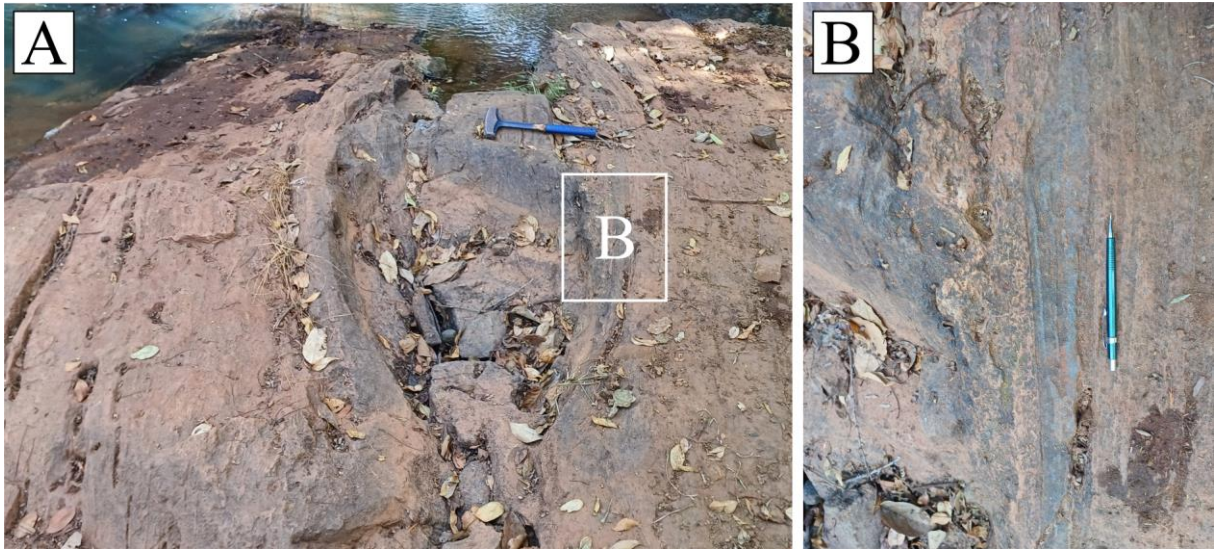


**Figura 17. Afloramento de rochas metabásicas do ponto TFLB\_24.**

Os pontos TFLB\_09 e TFLB\_10 (Figura 18 e 19) estão localizados dentro da delimitação do Corpo Capelinha dos Gomes, em pontos de predição e em zonas de moderada a altíssima probabilidade (60-100%) de ocorrência da classe Corpos M-UM. No ponto TFLB\_09 foram observados quartzo-dioritos, rochas mesocráticas a melanocráticas de granulação média com presença de foliação ígnea subvertical (Figura 18A e Figura 18B). No ponto TFLB\_10 foram observados hornblenda-gabros, rochas compostas essencialmente por hornblenda, biotita e plagioclásio. Essencialmente nesse último ponto, ocorrem veios félsicos centimétricos e enclaves.



**Figura 18. (A) Matakão de composição gabroica na porção centro-norte do Corpo Capelinha dos Gomes (Ponto: TFLB\_09). (B) Detalhe da parte superior do matakão da imagem “(A)”, exibindo foliação formada pela orientação mineral (135/80), possível feição de borda.**



**Figura 19. (A) Enclave máfico em rochas intermediárias com foliação subverticalizada (123/81) (Ponto TFLB\_10). (B) Detalhe para a borda do enclave evidenciando a segregação de bandas melanocráticas e leucocráticas.**

O ponto TFLB\_11 (Figura 20) está localizado dentro da delimitação do Corpo Fazenda Paraíso, em pontos de predição e em zonas de moderada a alta probabilidade (40-80%) de ocorrência da classe Corpos M-UM. Nestes pontos foram observados gabronoritos, rochas de granulação média a grossa, magnetismo moderado e compostas principalmente por piroxênio e plagioclásio, alternando sua granulação em diferentes pontos do afloramento.



**Figura 20. Encosta de morro composta por blocos desarticulados de gabronoritos nas intermediações do Corpo Fazenda Paraíso (Ponto TFLB\_11).**

Os demais pontos (TFLB\_13, TFLB\_14, TFLB\_16, TFLB\_17 e TFLB\_26) estão localizados fora das delimitações oficiais dos corpos intrusivos presentes nesta região, nas proximidades de acúmulos formados por pontos preditivos.

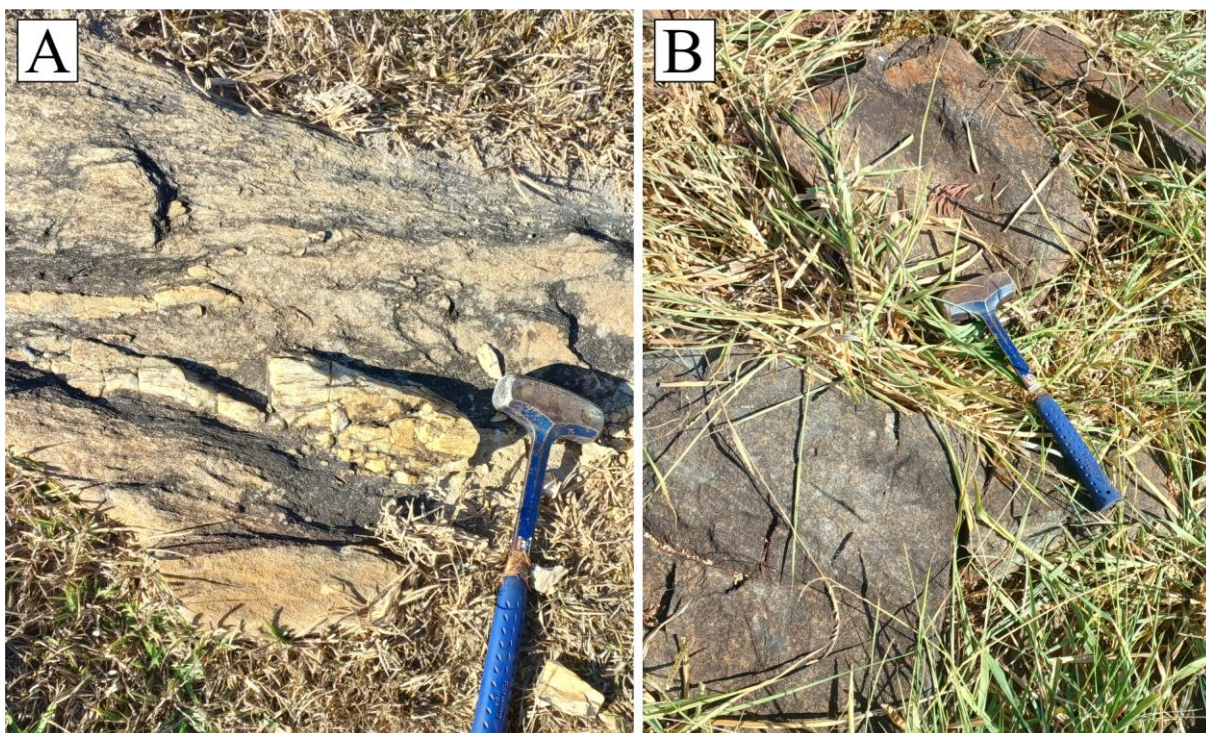
O ponto TFLB\_13 (Figura 21) está localizado fora da delimitação de qualquer corpo intrusivo, oficialmente no contato entre as unidades Metagranito Campestre (NP2camg) e Ortogranulítica Básica (NP1aitob) do CAI. Esse acúmulo de predições possui entre baixa a razoável probabilidade (40-60%) de ocorrência da classe Corpos M-UM. Neste ponto há um afloramento decamétrico no leito do Córrego do Buraco de metabasitos, holomelanocráticos, afaníticos, com foliação incipiente. O litotipo encontrado nesse ponto (Figura 21B), assemelha-se aos metabasitos do Corpo Águas Claras de Goiás.



**Figura 21. (A) Afloramento de metabasitos nas intermediações do Corpo Águas Claras de Goiás (Ponto: TFLB\_13). (B) Metabasito (Ponto: TFLB\_13).**

O ponto TFLB\_16 (Figura 22A e Figura 22B) também está localizado fora da delimitação de qualquer corpo intrusivo, oficialmente alocado na unidade Metagranito Campestre (NP2camg) do CAI. Esse acúmulo de predições possui entre baixa a razoável probabilidade (40-60%) de ocorrência da classe Corpos M-UM. Neste ponto há um afloramento de micaxistos com dobras intrafoliares e presença de veios quartzosos (Figura 22A). Nesse mesmo ponto, a

aproximadamente 50m de distância dos micaxistos também há um conjunto de blocos desarticulados de metadiorito.



**Figura 22. (A) Micaxistos em processo de gnaissificação (Ponto TFLB\_16). (B) Blocos de metadiorito (Ponto TFLB\_16).**

O ponto TFLB\_14 (Figura 23) está localizado fora da delimitação de qualquer corpo intrusivo, oficialmente locado no contato entre as unidades Ortogranulítica (NP1aito) e Ortogranulítica Básica (NP1aitob) do CAI. Esse acúmulo de predições possui entre razoável a alta probabilidade (60-80%) de ocorrência da classe Corpos M-UM. Este ponto está localizado em um alto topográfico, o qual aparenta ser composto majoritariamente por hornblenda-metagabronoritos com deformação incipiente, que gera uma leve orientação mineral. Composto essencialmente por piroxênio, anfibólio (hornblenda) e plagioclásio. As rochas desse ponto encontram-se em um grau moderado de alteração, gerando uma coloração esbranquiçada nas superfícies dos planos de fraturamento (Figura 23 C).



**Figura 23. (A) Visão ampla do afloramento do ponto TFLB\_14. (B) Blocos rolados de metagabronorito (TFLB\_14). (C) Detalhe para o metagabronorito do ponto TFLB\_14 com foliação evidente.**

O ponto TFLB\_17 (Figuras 24 e 25) está localizado fora da delimitação de qualquer corpo intrusivo, oficialmente alocado na unidade Ortogranulítica (NP1aito) do CAI. Esse acúmulo de predições possui entre razoável e alta probabilidade (60-80%) de ocorrência da classe Corpos M-UM. Este ponto está localizado em um alto topográfico, o qual aparenta ser composto majoritariamente pelos litotipos encontrados. Trata-se de um ortognaisse básico milonitizado.



**Figura 24. Ortognaisse básico milonitizado (Ponto TFLB\_17) com detalhe para a foliação milonítica da amostra coletada no ponto.**



**Figura 25. (A) Blocos desarticulados de ortognaisse migmatizado (Ponto TFLB\_17). (B) Detalhe para uma banda melanocrática dos ortognaisses do Ponto TFLB\_17.**

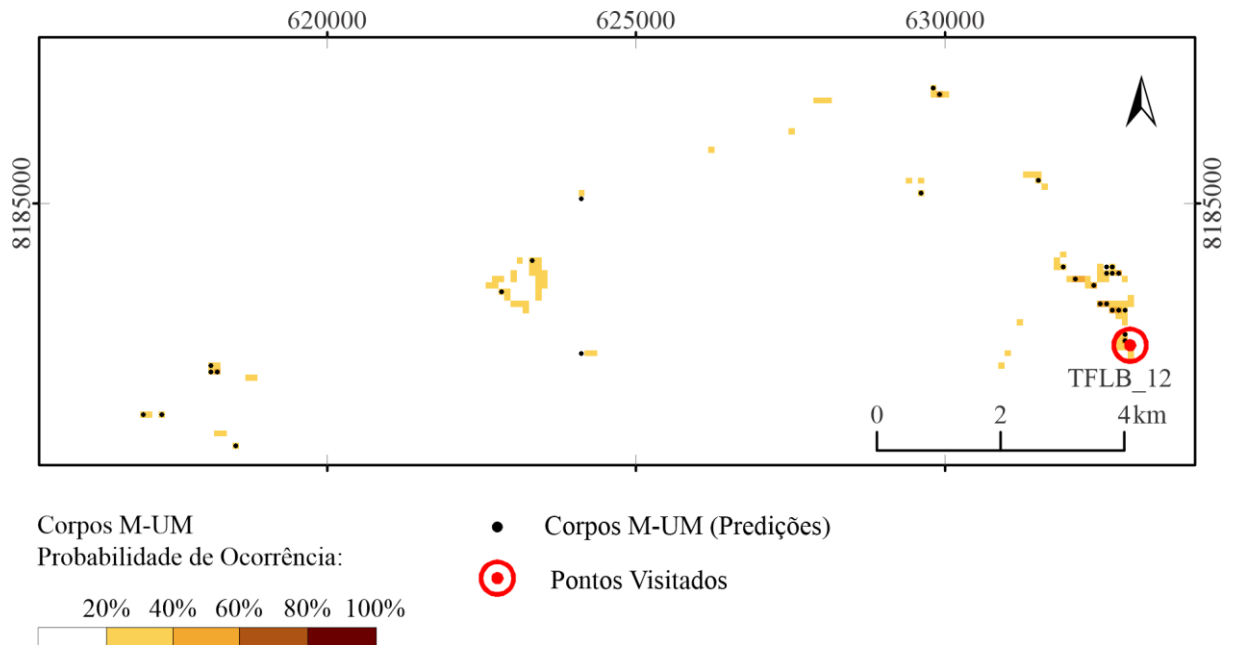
O ponto TFLB\_26 (Figura 26) está localizado fora da delimitação de qualquer corpo intrusivo, oficialmente alocado na unidade Ortogranulítica (NP1aito) do CAI, próximo a um acúmulo de predições. As proximidades desse ponto possuem entre baixa a razoável probabilidade (40-60%) de ocorrência da classe Corpos M-UM. Neste ponto há um afloramento composto por lajes e matacões às margens do Córrego Capoeirão. As rochas observadas foram hornblendagabros de granulação média, compostos essencialmente por piroxênio, hornblenda e plagioclásio.



**Figura 26.** Afloramento de hornblenda-gabros do ponto TFLB\_26.

### **5.3.5 REGIÃO 5: SE DA ÁREA DE ESTUDO - PREDIÇÕES ISOLADAS**

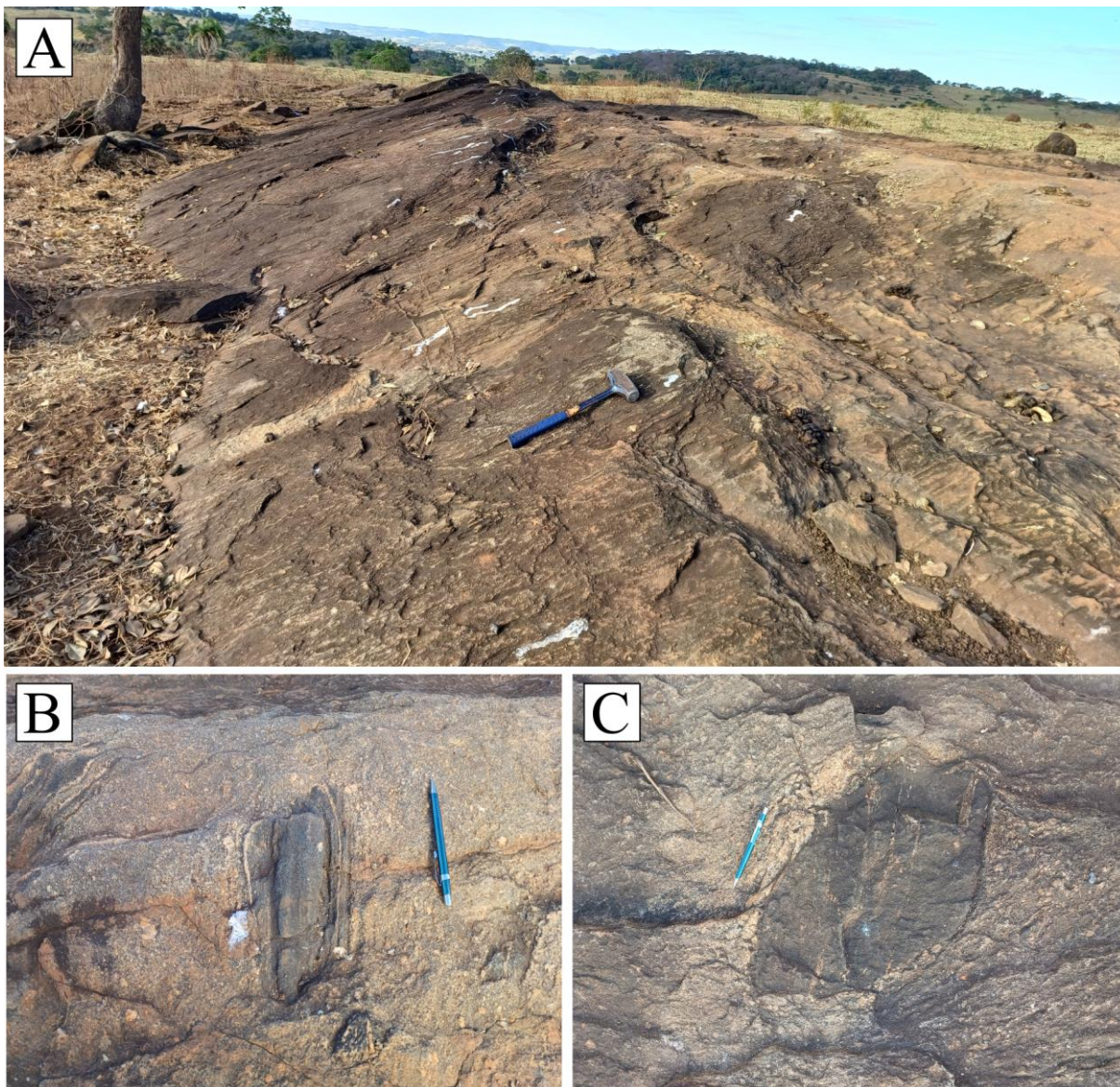
Na região sudeste da área de estudo (Figura 27), foram identificadas poucas predições isoladas que indicam a presença de corpos máficos-ultramáficos. Embora essas predições estejam dispersas, a validação em campo revelou que o principal acúmulo formado, extremo leste da região, corresponde a litologias com associações máficas.



**Figura 27. Mapa preditivo com enfoque nas predições isoladas na região sudeste da área de estudo, indicando a probabilidade de ocorrência da classe de Corpos Máficos-Ultramáficos na área de estudo.**

O ponto TFLB\_12 está localizado fora da delimitação de qualquer corpo intrusivo, oficialmente na unidade Metagranito Campestre (NP2camg). Esse acúmulo de predições possui entre baixa a razoável probabilidade (40-60%) de ocorrência da classe Corpos M-UM. Neste ponto há um afloramento decamétrico de ortognaisses com enclaves máficos (Figura 28). As rochas variam entre os índices de cor mesocrático e melanocrático, com separação de bandas félsicas e máficas, com granulação fina a média. Os enclaves possuem composição diorítica.

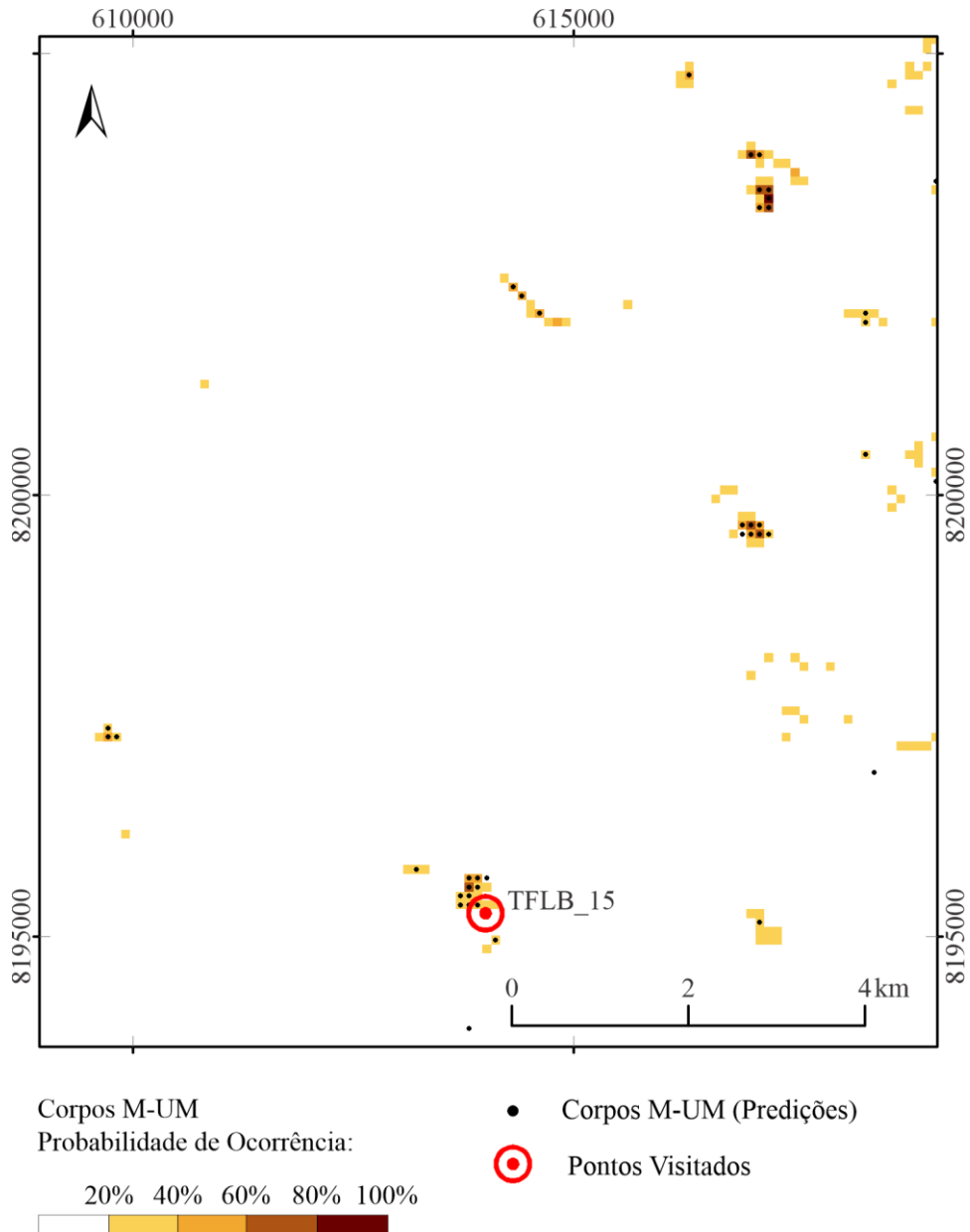
A dispersão dos pontos preditivos sugere uma complexidade geológica elevada para o modelo, onde pequenas porções de composição semelhante a de corpos máficos, podem estar presentes de forma disseminada formadas por eventos menores. A verificação em campo dessas predições confirma a utilidade do modelo *Random Forest* em identificar corretamente áreas de interesse, mesmo em regiões com predições menos concentradas.



**Figura 28. (A) Afloramento decamétrico composto por ortognaisses (Ponto TFLB\_12). (B e C) Detalhe para os enclaves máficos contidos nos ortognaisses do ponto TFLB\_12.**

### **5.3.6 REGIÃO 6: CENTRO DA ÁREA DE ESTUDO - PREDIÇÕES ISOLADAS**

O centro da área de estudo (Figura 29) apresentou várias predições isoladas, semelhantes às observadas na região sudeste. Por facilidade de acesso, um único ponto (TFLB\_15) foi visitado, e validou que os pontos de predição coincidem com áreas onde foram encontradas rochas de composição máfica.



**Figura 29. Mapa preditivo com enfoque nas predições isoladas no centro da área de estudo, destacando a identificação de corpos máficos-ultramáficos pelo modelo Random Forest.**

O ponto TFLB\_15, assim como o ponto TFLB\_12, está localizado fora da delimitação de qualquer corpo intrusivo, e oficialmente na unidade Ortogranulítica Ácida do CAI (NP1aitoa). Esse acúmulo de predições possui entre razoável a alta probabilidade (60-80%) de ocorrência da classe Corpos M-UM. Neste ponto afloram blocos de hornblenda-gabronoritos (Figura 30) de granulação média a grossa, magnetismo moderado, compostos principalmente por piroxênio, plagioclásio, hornblenda, biotita e quartzo (<5%).

A confirmação dessas previsões em campo reforça a confiabilidade do modelo e a sua aplicabilidade em diferentes contextos geológicos, uma vez que, pode identificar dentro da unidade Ortogranulítica do CAI as porções com características semelhantes às dos Corpos Intrusivos M-UM.



Figura 30. (A) Blocos de hornblenda-gabronorito desarticulados (Ponto TFLB\_15).

## 6. CONSIDERAÇÕES FINAIS

As técnicas de *Machine Learning* aplicadas neste estudo demonstraram ser métodos eficientes e de baixo custo para a identificação de alvos preliminares e o refino de estratégias de campo. Estes métodos podem ser aplicados em outras regiões ou contextos, proporcionando uma ferramenta valiosa para a geologia aplicada. A utilidade dos mapas preditivos gerados neste estudo é multifacetada e pode ser explorada em diversas etapas do trabalho geológico. Como material de pré-campo, os mapas preditivos podem servir como um recurso valioso durante a fase de planejamento das campanhas de campo. Eles fornecem uma visão inicial das áreas de interesse, permitindo que os geólogos otimizem suas rotas de amostragem e concentrem seus esforços nas zonas com maior probabilidade de conter as litologias de interesse. Isso não só economiza tempo, mas também reduz os custos operacionais associados às campanhas de campo.

A integração de dados geofísicos e de sensoriamento remoto com os modelos de *Machine Learning* permite inferir contatos litológicos de forma mais precisa. Os produtos geofísicos, como magnetometria e gamaespectroscopia, combinados com dados de sensoriamento remoto, como imagens de satélite e modelos digitais de elevação, ajudam a identificar anomalias e padrões que indicam a presença de diferentes litologias. Os mapas preditivos auxiliam na interpretação desses dados, fornecendo uma base mais robusta para a inferência de contatos litológicos. Além disso, a utilização de mapas preditivos facilita a análise integrada de diferentes tipos de dados geológicos, incluindo a correlação de informações litológicas, geofísicas e topográficas, resultando em uma compreensão mais completa da geologia da área de estudo. Esta abordagem integrada é essencial para a prospecção mineral e para a identificação de novos alvos exploratórios.

Para estudos futuros, recomenda-se a aplicação de técnicas de aprendizado profundo (*deep learning*) para melhorar ainda mais a precisão das previsões. A utilização de dados de afloramento para o treinamento de modelos de ML pode ser uma alternativa promissora para refinar a identificação de determinadas litologias. Além disso, a expansão do estudo para outras áreas ajudará a validar ainda mais a eficácia das técnicas de *Machine Learning* na geologia, como a identificação de áreas de risco potencial, cavernas e potencial metalogenético. Este estudo contribui significativamente para o conhecimento da geologia da região de Goiás, oferecendo uma abordagem inovadora para a identificação de corpos máficos-ultramáficos. A aplicação de *Machine Learning* na prospecção mineral mostra-se promissora e pode revolucionar a forma como estudos geológicos são conduzidos.

O uso de algoritmos de *Machine Learning*, como o *Random Forest*, demonstrou ser uma abordagem eficaz para a identificação e classificação de litologias máficas e ultramáficas. A validação em campo confirmou a precisão das previsões, destacando a robustez do modelo aplicado. Estudos futuros poderão se beneficiar da integração de técnicas avançadas de aprendizado profundo e da ampliação do escopo geográfico para validar ainda mais a eficácia dessas abordagens.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

AGRESTI, Alan. Categorical data analysis. John Wiley & Sons, 2012.

ALMEIDA, Fernando Flávio Marques. Origem e evolução da plataforma brasileira. DNPM, 1967.

ALMEIDA, Fernando Flávio Marques et al. Brazilian structural provinces: an introduction. *Earth-Science Reviews*, v. 17, n. 1-2, p. 1-29, 1981.

BERGSTRA, James; BENGIO, Yoshua. Random search for hyper-parameter optimization. *Journal of machine learning research*, v. 13, n. 2, 2012.

BISHOP, Christopher M.; NASRABADI, Nasser M. Pattern recognition and machine learning. New York: springer, 2006.

BREIMAN, Leo. Random forests. *Machine learning*, v. 45, p. 5-32, 2001.

BREIMAN, Leo. Classification and regression trees. Routledge, 2017.

CARNEIRO, Jônatas; FUCK, Reinhardt; DANTAS, Elton Luiz. Arenópolis sequence, evolution of a marginal basin in the Neoproterozoic Goiás magmatic arc, central Brazil. *Journal of South American Earth Sciences*, v. 106, p. 103033, 2021.

CARRANZA, Emmanuel John M.; LABORTE, Alice G. Random forest predictive modeling of mineral prospectivity with small number of prospects and data with missing values in Abra (Philippines). *Computers & Geosciences*, v. 74, p. 60-70, 2015.

COSTA, Iago Sousa Lima; TAVARES, Felipe Mattos; DE OLIVEIRA, Junny Kyle Mastop. Predictive lithological mapping through machine learning methods: a case study in the Cinzento Lineament, Carajás Province, Brazil. *Journal of the Geological Survey of Brazil*, v. 2, n. 1, p. 26-36, 2019.

COSTA, Iago Sousa Lima et al. Uranium anomalies detection through Random Forest regression. *Exploration Geophysics*, v. 51, n. 5, p. 555-569, 2020.

CRACKNELL, M. J.; READING, A. M.; MCNEILL, A. W. Mapping geology and volcanic-hosted massive sulfide alteration in the Hellyer–Mt Charter region, Tasmania, using Random Forests™ and Self-Organising Maps. *Australian Journal of Earth Sciences*, v. 61, n. 2, p. 287-304, 2014.

CPRM – Serviço Geológico do Brasil. Levantamento aerogeofísico do estado de Goiás-1ª etapa: arco magmático de Arenópolis: Complexo Anápolis-Itauçu, seqüência vulcano-sedimentar de Juscelândia: relatório final do levantamento e processamento dos dados magnetométricos e gamaespectrométricos. Lasa Engenharia e Prospecções, 2004.

DARDENNE, M. A. The Brasília Fold Belt. In: CORDANI, U. G.; MILANI, E. J.; THOMAZ FILHO, A.; CAMPOS, D. A. (Eds.). Tectonic Evolution of South America. Rio de Janeiro: 31st International Geological Congress, p. 231-263, 2000.

FUCK, R. A. A Faixa Brasília e a compartimentação tectônica na Província Tocantins. SBG, Simpósio de Geologia do Centro-Oeste, v. 4, p. 184-187, 1994.

FUCK, R. A.; PIMENTEL, M. M.; SILVA, L. J. H. D. Compartimentação tectônica na porção oriental da Província Tocantins. In: Congresso Brasileiro de Geologia. SBG Balneário Camboriú, 1994. p. 215-216.

GHOLAMY, Afshin; KREINOVICH, Vladik; KOSHELEVA, Olga. Why 70/30 or 80/20 relation between training and testing sets: A pedagogical explanation. Int. J. Intell. Technol. Appl. Stat, v. 11, n. 2, p. 105-111, 2018.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. Deep learning. MIT press, 2016.

HASTIE, Trevor et al. The elements of statistical learning: data mining, inference, and prediction. New York: springer, 2009.

KOHAVI, Ron; PROVOST, Foster. Glossary of terms. Machine Learning—Special Issue on Applications of Machine Learning and the Knowledge Discovery Process. Machine Learning, 30, 271-274, 1998.

KUHN, Stephen; CRACKNELL, Matthew J.; READING, Anya M. Lithologic mapping using Random Forests applied to geophysical and remote-sensing data: A demonstration study from the Eastern Goldfields of Australia. Geophysics, v. 83, n. 4, p. B183-B193, 2018.

KUHN, Stephen; CRACKNELL, Matthew J.; READING, Anya M. Lithological mapping in the Central African Copper Belt using Random Forests and clustering: Strategies for optimised results. Ore Geology Reviews, v. 112, p. 103015, 2019.

LACERDA FILHO, Joffre Valmório de et al. Projeto geologia e metalogenia da porção oeste de Goiás. CPRM, 2021.

LAUX, Jorge Henrique. Evolução do Arco Magmático de Goiás com base em dados geocronológicos U-Pb e Sm-Nd. 2004. Tese de Doutorado.

LAUX, Jorge Henrique et al. Mafic magmatism associated with the Goiás magmatic arc in the Anicuns region, Goiás, central Brazil: Sm–Nd isotopes and new ID-TIMS and SHIMP U–Pb data. Journal of South American Earth Sciences, v. 16, n. 7, p. 599-614, 2004.

LORENA, Ana Carolina; DE CARVALHO, André CPLF. Uma introdução às support vector machines. Revista de Informática Teórica e Aplicada, v. 14, n. 2, p. 43-67, 2007.

MITCHELL, Tom M.; MITCHELL, Tom M. Machine learning. New York: McGraw-hill, 1997.

PIMENTEL, Márcio M. et al. The Mara Rosa arch in the Tocantins province: Further evidence for neoproterozoic crustal accretion in Central Brazil. *Precambrian Research*, v. 81, n. 3-4, p. 299-310, 1997.

PIMENTEL, M. M.; JOST, Hardy; FUCK, Reinhardt A. O embasamento da Faixa Brasília e o Arco Magmático de Goiás. *Geologia do continente Sul-americano: evolução da obra de Fernando Flávio Marques de Almeida*, v. 1, p. 355-368, 2004.

POWERS, David MW. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. arXiv preprint arXiv:2010.16061, 2020.

QUINLAN, J.. Ross . Induction of decision trees. *Machine learning*, v. 1, p. 81-106, 1986.

QUINLAN, J. Ross. Generating production rules from decision trees. In: *ijcai*. 1987. p. 304-307.

RODRIGUEZ-GALIANO, Victor Francisco et al. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS journal of photogrammetry and remote sensing*, v. 67, p. 93-104, 2012.

SILVA, Jonas Mota. O Complexo Máfico-ultramáfico Acamadado de Americano do Brasil e sua mineralização de Ni-Cu-Co. 2009. Dissertação de Mestrado em Geologia, Universidade de Brasília, Brasília, 2009.

SILVA, Jonas Mota et al. Geology, petrology and geochemistry of the “Americano do Brasil” layered intrusion, central Brazil, and its Ni–Cu sulfide deposits. *Mineralium Deposita*, v. 46, p. 57-90, 2011.

STOICA, Petre; BABU, Prabhu. Pearson-Matthews correlation coefficients for binary and multinary classification and hypothesis testing. arXiv preprint arXiv:2305.05974, 2023.

STRECKEISEN, Albert. To each plutonic rock its proper name. *Earth-science reviews*, v. 12, n. 1, p. 1-33, 1976.

TABOSA, Luiz Fernando Cursino et al. Utilizing Random Forest for Identifying Mafic and Ultramafic Rocks in the Gameleira Suite of the Paleoproterozoic Archean Basement in the Brasília Belt, Brazil.

VALERIANO, C. de M. et al. Tectonic evolution of the Brasília Belt, Central Brazil, and early assembly of Gondwana. *Geological Society, London, Special Publications*, v. 294, n. 1, p. 197-210, 2008.

VIANA, Maria Das Graças et al. O arco magmático de Mara Rosa, Goiás: dados geoquímicos e geocronológicos e suas implicações regionais. *Brazilian Journal of Geology*, v. 25, n. 2, p. 111-123, 1995.

WINGE, MANFREDO. Evolução dos terrenos granulíticos da Província Estrutural Tocantins, Brasil central. Tese de Doutorado, Universidade de Brasília, 1995.

## **7.1 REFERÊNCIAS DE *SOFTWARES* UTILIZADOS**

ArcMap. Versão 10.8. Esri. Redlands, CA: 2021.

CorelDRAW. Versão 2022. Corel Corporation. Ottawa, ON: 2021.

QGIS. Versão 3.36.3. QGIS.org, 2024.

Microsoft Word. Versão 2023. Microsoft Corporation, 2023.

Microsoft Excel. Versão 2023. Microsoft Corporation, 2023.

Orange. Versão 3.33. University of Ljubljana, 2023.

Oasis Montaj. Versão 9.10. Seequent, 2023.

Python. Versão 3.11.4. Python Software Foundation, 2023.

## **7.2 REFERÊNCIAS DE PLATAFORMAS UTILIZADOS**

Earth Explorer. Disponível em: <https://earthexplorer.usgs.gov/>. Acesso em: 8 jul. 2024.

RiGeo. Repositório Institucional de Geociências. Disponível em: <https://rigeo.cprm.gov.br/>. Acesso em: 8 jul. 2024.



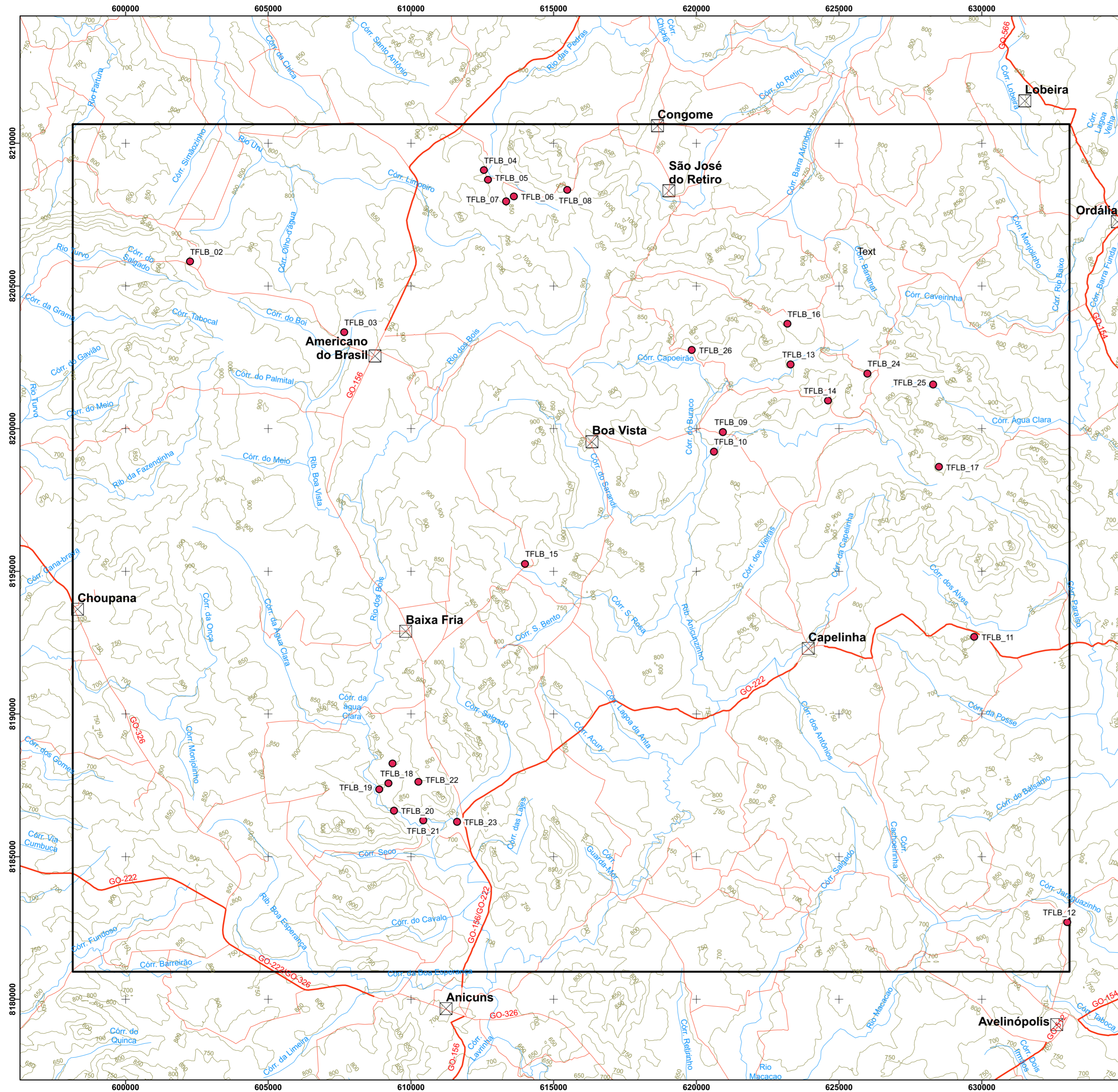
## APÊNDICE B: TABELA DE PONTOS

Ponto	X	Y	Elevação	Tipo	Litotipo	Descrição
TFLB_01	609355	8188264	793	Solo	Solo argiloso avermelhado.	Vegetação densa, não foi possível observar afloramentos. Solo argiloso avermelhado. Resposta de subsuperfície (?)
TFLB_02	602258	8205857	911	Afloramento	Piroxenito	Afloramento decamétrico, composto por blocos e matacões semi-articulados, moderadamente intemperizados. Rocha melanocrática a holomelanocrática, fanerítica, granulação grossa a muito grossa, inequigranular porfirítica. Rocha composta por px, anf e sulfetos.
TFLB_03	607664	8203376	879	Solo	Solo argilo-siltoso marrom claro	Solo argilo-siltoso marrom claro, não foi possível observar afloramentos. Área de plantio.
TFLB_04	612560	8209057	845	Blocos Autóctones	Diorito, Gabro	Blocos e matacões desarticulados moderadamente intemperizados. A rocha é mesocrática a melanocrática, de granulação média, inequigranular, composta em suma por px-plg-bt-qtz e sulfetos.
TFLB_05	612698	8208718	845	Afloramento	Piroxenito	Piroxenito (websterito?) de granulação média a grossa, inequigranular porfirítica, composta essencialmente por piroxênios. Pouco plg e minerais acessórios.
TFLB_06	613608	8208134	882	Blocos Autóctones	Quartzo-Diorito	Blocos submétricos a métricos desarticulados. Quartzo-diorito, maciço, inequigranular de granulação média. Composto por px-plg-hbl-bt-qtz.
TFLB_07	613335	8207969	855	Afloramento	Xisto máfico	Afloramento em laje de Xisto máfico(?). Chl-tlc-anf-ms-xisto, mesocrático, bem intemperizado, magnetismo incipiente.
TFLB_08	615479	8208363	945	Afloramento	Grt-gabro	Blocos e matacões gabróicos, maciços, inequigranular porfiríticos, de granulação média a grossa, magnetismo moderado e compostos

<b>Ponto</b>	<b>X</b>	<b>Y</b>	<b>Elevação</b>	<b>Tipo</b>	<b>Litotipo</b>	<b>Descrição</b>
						principalmente por px-plg-grt e sulfetos em menor quantidade.
TFLB_09	620931	8199878	770	Afloramento	Quartzo-diorito	Afloramento no leito da drenagem. Rochas mesocráticas a melanocráticas de granulação média com presença de foliação ígnea subvertical (feição de borda?)
TFLB_10	620622	8199193	743	Afloramento	Hornblenda-gabro	Rochas compostas essencialmente por hornblenda, biotita e plagioclásio
TFLB_11	629742	8192697	771	Blocos Autóctones	Gabronorito	Rochas de granulação média a grossa, magnetismo moderado e compostas principalmente por px e plg,
TFLB_12	632998	8182694	751	Afloramento	Ortognaisse	Ortognaisse com enclaves máficos. Separação de bandas félsicas e máficas, com granulação fina a média. Os enclaves possuem composição diorítica.
TFLB_13	623302	8202250	803	Afloramento	Metabasito	Aafloramento decamétrico no leito do Córrego do Buraco. Metabasitos, holomelanocráticos, afaníticos, com foliação incipiente e indícios de serpentinização.
TFLB_14	624619	8200980	824	Blocos Autóctones	Hornblenda metagabronorito	Hornblenda-metagabronoritos com deformação incipiente, que gera uma leve orientação mineral. Composto essencialmente por piroxênio, anfibólio (hornblenda) e plagioclásio.
TFLB_15	613994	8195258	794	Blocos Autóctones	Gabronorito	Blocos desarticulados de hornblenda-gabronoritos de granulação média a grossa, magnetismo moderado, compostos principalmente por px, plg, hbl, bt e qtz (<5%).
TFLB_16	623191	8203674	862	Afloramento	Mica-xisto, Tonalito	Afloramento de micaxistos com dobras intrafoliais e presença de veios quartzosos. No entorno há blocos desarticulados de metadiorito.
TFLB_17	628500	8198656	928	Afloramento	Milonito	Afloramento no topo do morro, próximo a Fazenda Brasília. Lajes, blocos e matações. Ortognaisse milonitizado, visível distinção entre

Ponto	X	Y	Elevação	Tipo	Litotipo	Descrição
						bandas melanocráticas (hbl+bt) e leucocráticas(plg+qtz).
TFLB_18	609213	8187569	760	Blocos Autóctones	Granodiorito, Diorito	Rocha classificada como granítica/granodiorítica, de coloração cinza, moderadamente alterada, fracamente magnética. Composição mineralógica <10% de plg, <5% de bt, mais de 20% de qtz. Ocorre epidoto como mineral de alteração. Textura fanerítica, granulação fina a média. Ocorrem veios de quartzo e diques aplíticos com direção NW-SE.
TFLB_19	608892	8187354	750	Afloramento	Granodiorito, Diorito	Ocorrem 3 litotipos no afloramento: Granito/granodiorito; diques de diabásio métricos; veios aplíticos centimétricos. É visível a mistura física de magmas.
TFLB_20	609405	8186611	732	Afloramento	Gabro	Aqui afloram gabros, granulação média, compostos por plg e piroxênios. Ocorrem também diques de diabásio com direção Az. 045. Pelo menos 4 famílias de fraturas nesse ponto.
TFLB_21	610434	8186272	702	Afloramento	Gabro	Afloramento localizado as margens do Rio dos Bois disposto na forma de lajedo.
TFLB_22	610262	8187617	826	Blocos Autóctones	Gabro	Afloramento de blocos in situ, transição da composição anterior para rochas gabróicas. Blocos métricos a centimétricos em um alto topográfico. Possui leve magnetismo. Variação faciológica é visível.
TFLB_23	611612	8186220	721	Blocos Autóctones	Gabro	Afloramento de blocos desarticulados no topo do morro à W da estrada GO-222 diametralmente à GO-156. Rocha de composição diorítica, granulação média à grossa.
TFLB_24	625998	8201932	783	Afloramento	Metagabro,	Leito da drenagem. Rochas metabásicas holomelanocráticas afaníticas e metagabros de granulação fina a média, compostos principalmente por piroxênio e plagioclásio.

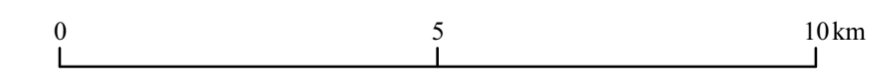
<b>Ponto</b>	<b>X</b>	<b>Y</b>	<b>Elevação</b>	<b>Tipo</b>	<b>Litotipo</b>	<b>Descrição</b>
TFLB_25	628303	8201547	819	Afloramento	Metabasito, Gabros	Leito da drenagem. Rochas metabásicas holomelanocráticas afaníticas e metagabros de granulação fina a média, compostos principalmente por piroxênio e plagioclásio.
TFLB_26	619843	8202754	775	Afloramento	Hornblenda-Gabro	Hornblenda-gabros de granulação média, compostos essencialmente por piroxênio, hornblenda e plagioclásio.

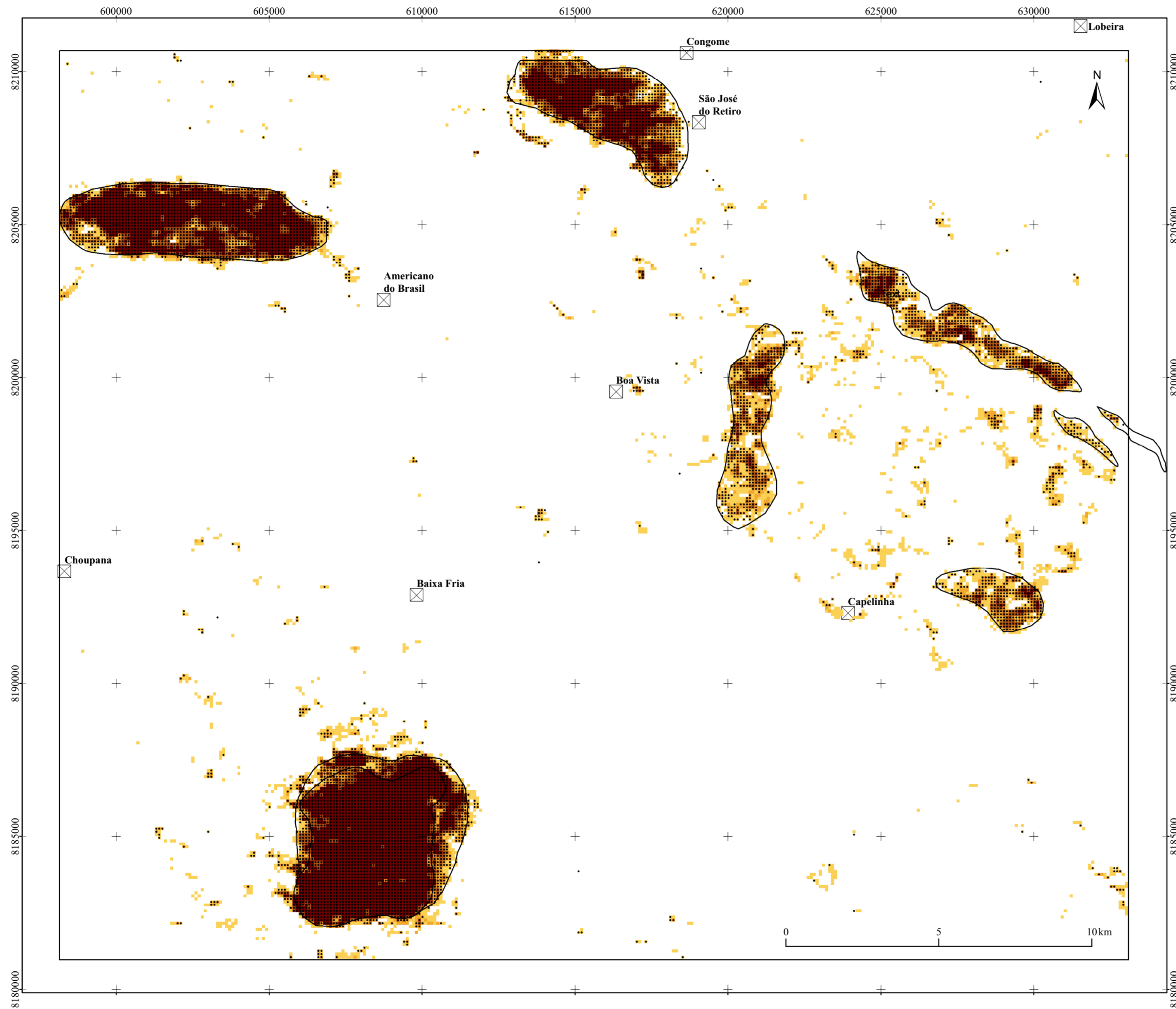


**APÊNDICE C:**  
**Mapa de Pontos**  
 PROJEÇÃO UNIVERSAL TRANSVERSA DE MERCATOR  
 Datum horizontal: SIRGAS 2000  
 Escala 1:100000  
 Junho 2024

**Convenções Cartográficas**

- Pontos Visitados
- Área de Estudo
- Rodovias
- Drenagens
- Cidades
- Curvas de nível (50m)





**APÊNDICE D:  
Mapa Preditivo de Corpos Máficos-Ultramáficos**

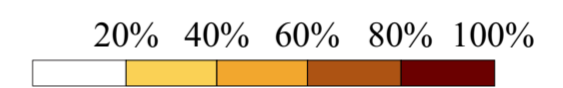
PROJEÇÃO UNIVERSAL TRANSVERSA DE MERCATOR

Datum horizontal: SIRGAS 2000

Escala 1:100000

Junho 2024

Corpos M-UM  
Probabilidade de Ocorrência:



□ Corpos M-UM (Lacerda Filho et. al., 2021)

• Corpos M-UM (Predições)

⊠ Cidades